

541.750



(10) Internationale Veröffentlichungsnummer
WO 2004/063359 A2

(72) Erfinder; und
(75) Erfinder/Anmelder (nur für US): MATUSCHEK, Markus [DE/DE]; Karolinenstr. 5, 69469 Weinheim (DE). KLEIN, Daniela [DE/DE]; M 7, 2, 68161 Mannheim (DE). HEINEKAMP, Thorsten [DE/DE]; Alte Ziegelei 1B, 30419 Hannover (DE). SCHMIDT, Andre [DE/DE]; Magdeburger Str. 11, 31832 Springe (DE). BRAKHAGE, Axel [DE/DE]; Schneiderberg 58, 30167 Hannover (DE). ACHATZ, Brigitte [DE/DE]; Windeckstr. 26, 68163 Mannheim (DE).

(75) **Erfinder/Anmelder (nur für US): MATUSCHEK, Markus [DE/DE];** Karolinenstr. 5, 69469 Weinheim (DE). **KLEIN, Daniela [DE/DE];** M 7, 2, 68161 Mannheim (DE). **HEINEKAMP, Thorsten [DE/DE];** Alte Ziegelei 1B, 30419 Hannover (DE). **SCHMIDT, Andre [DE/DE];** Magdeburger Str. 11, 31832 Springe (DE). **BRAKHAGE, Axel [DE/DE];** Schneiderberg 58, 30167 Hannover (DE). **ACHATZ, Brigitte [DE/DE];** Windeckstr. 26, 68163 Mannheim (DE).

(74) Anwalt: FITZNER, Uwe; Lintorfer Str. 10, 40878 Ratingen (DE).

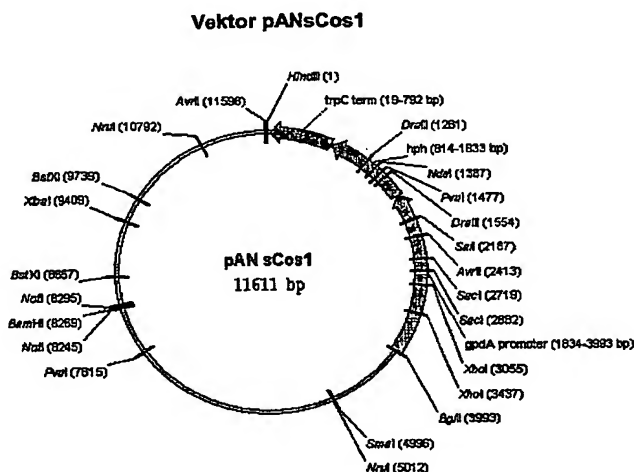
(81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,

[Fortsetzung auf der nächsten Seite]

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,

(54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG VON CAROTINOIDEN ODER DEREN VORSTUFEN MITTELS GENTECHNISCH VERÄNDERTER ORGANISMEN DER GATTUNG *BLAKESLEA*, MIT DEM VERFAHREN HERGESTELLTE CAROTINOIDE ODER DEREN VORSTUFEN UND DEREN VERWENDUNG



[Fortsetzung auf der nächsten Seite]

WO 2004/063359 A2



FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— *ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts*

(84) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK,

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) **Zusammenfassung:** Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderter Organismen der Gattung *Blakeslea* umfassend (i) (i) Transformation mindestens einer der Zellen, (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und (ix) Selektion und Vermehrung der gentechnisch veränderten Zelle oder Zellen, (x) Kultivierung der gentechnisch veränderten Zellen, (xi) Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe; nach dem Verfahren hergestellte Carotinoide oder deren Vorstufen und deren Verwendung.

**Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen
mittels gentechnisch veränderter Organismen der Gattung Blakes-
lea, mit dem Verfahren hergestellte Carotinoide oder deren Vorstufen
und deren Verwendung**

5

Die Erfindung betrifft ein Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderten Organismen der Gattung Blakeslea, mit dem Verfahren hergestellte Carotinoide oder deren Vorstufen und deren Verwendung und Bereitstellung, besonders als hoch-

10 reine Carotinoide, als Nahrungsmittel, enthaltend Carotinoideproduzierende Organismen und mindestens ein Carotinoid, insbesondere Tierfuttermittel, Tierfutterergänzungsmittel und Nahrungsergänzungsmittel, sowie die Verwendung der aus dem Verfahren erhältlichen Carotinoide zur

15 Herstellung von kosmetischen, pharmazeutischen, dermatologischen Zubereitungen, Nahrungsmitteln oder Nahrungsergänzungsmitteln.

Blakeslea trispora ist als Produktionsorganismus für β -Carotin (Ciegler, 1965, Adv Appl Microbiol. 7:1) und Lycopin bekannt (EP 1201762, EP 1184464, WO 03/038064).

20

Von Blakeslea trispora sind bisher verschiedene DNA-Sequenzen bekannt, insbesondere die DNA-Sequenz, die für die Gene der Carotinoidbiosynthese von Geranylgeranylpyrophosphat bis β -Carotin codiert (WO 03/027293).

25

Insbesondere aufgrund der hohen Produktivität, die mit Blakeslea in der Produktion von Lycopin und β -Carotin erreicht werden, bietet sich dieser Organismus zur fermentativen Herstellung von Carotinoiden an.

30

Es ist auch von Interesse die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Her-

stellung weiterer Carotinoide, wie z. B. Xanthophylle zu ermöglichen, die von Blakeslea bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

5 Carotinoide werden Futtermitteln, Nahrungsmitteln, Nahrungsergänzungsmitteln, Kosmetika und Arzneimitteln zugesetzt. Die Carotinoide dienen vor allem als Pigmente zur Färbung. Daneben werden die antioxidative Wirkung der Carotinoide und andere Eigenschaften dieser Substanzen genutzt. Man unterteilt die Carotinoide in die reinen Kohlenwasserstoffe,
10 die Carotine und die sauerstoffhaltigen Kohlenwasserstoffe, die Xanthophylle. Xanthophylle wie Canthaxanthin und Astaxanthin werden beispielsweise zur Pigmentierung von Hühnereiern und Fischen eingesetzt (Britton et al. 1998, Carotinoids, Vol 3, Biosynthesis and Metabolism). Die Carotine β -Carotin und Lycopin werden vor allem in der Humanernährung
15 eingesetzt. β -Carotin wird beispielsweise als Getränkefarbstoff verwendet. Lycopin hat eine krankheitsvorbeugende Wirkung (Argwal und Rao, 2000, CMAJ 163:739-744; Rao und Argwal 1999, Nutrition Research 19:305-323). Die farblose Carotinoidvorstufe Phytoen kommt vor allem für Anwendungen als Antioxidans in kosmetischen, pharmazeutischen oder
20 dermatologischen Zubereitungen in Frage.

Der überwiegende Teil der Carotinoide und deren Vorstufen, die als Zusatzstoffe für die oben genannten Anwendungen eingesetzt werden, wird durch chemische Synthese hergestellt. Die chemische Synthese ist technisch
25 sehr aufwendig und verursacht hohe Herstellkosten. Fermentative Verfahren sind demgegenüber technisch verhältnismäßig einfach und basieren auf kostengünstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden und deren Vorstufen können dann wirtschaftlich attraktiv und wettbewerbsfähig zur chemischen Synthese
30 sein, wenn die Produktivität der bisherigen fermentativen Verfahren gesteigert würde oder neue Carotinoide auf Basis der bekannten

würde oder neue Carotinoide auf Basis der bekannten Produktionsorganismen hergestellt werden könnten.

5 Hierzu ist eine gentechnische, d. h. gezielte genetische Veränderung von Blakeslea erforderlich. Insbesondere, wenn Xanthophylle produziert werden sollen, da diese Verbindungen natürlicherweise vom Wildtyp der Blakeslea nicht synthetisiert werden.

10 Z. B. zur Herstellung von Phytoen mittels Fermentation von Blakeslea trispora sind bisher zwei Methoden bekannt:

(i) Durch zufallsabhängige Mutagenese mit chemischen Agenzien wie MNNG können Mutanten erzeugt werden, in denen Phytoen nicht zu Lycopin und somit nicht weiter zu β -Carotin umgesetzt werden kann (Mehta und Cerdá-Olmedo, 1995, Appl. Microbiol. Biotechnol. 42:836-838).

20 (ii) Durch Zugabe von Inhibitoren des Enzyms Phytoendesaturase wie z.B. Diphenylamin und Zimtalkohol kann die weitere Umsetzung von Phytoen blockiert werden, so dass es sich anreichert (Cerdá-Olmedo, 1989, In: E. Vandamme, ed. Biotechnology of vitamin, growth factor and pigment production. London: Elsevier Applied Science, S. 27-42).

25 Die genannten Methoden zur Herstellung von Phytoen mit Blakeslea trispora weisen jedoch eine Reihe von Nachteilen auf.

30 Die zufallsabhängige Mutagenese betrifft in der Regel nicht nur die Gene der Carotinoidbiosynthese zur weiteren Umsetzung von Phytoen, sondern auch weitere wichtige Gene. Daher sind Wachstum und Syntheseleistung der Mutanten oft beeinträchtigt. Die Erzeugung z. B. von Phytoenüberpro-

duzenten durch zufallsabhängige Mutagenese von Lycopinüberproduzenten oder β -Carotinüberproduzenten ist daher entweder nicht oder nur mit großem experimentellem Aufwand zu erreichen. Die Zugabe von Inhibitoren verursacht eine Erhöhung der Produktionskosten und gegebenenfalls
5 eine Verunreinigung des Produktes. Daneben kann das Zellwachstum durch den Inhibitor beeinträchtigt werden, so dass die Produktion von Carotinoiden oder deren Vorstufen, insbesondere Phytoen eingeschränkt wird.

10 Durch eine gentechnische Veränderung könnten die oben genannten Nachteile der zufallsabhängigen Mutagenese und der Inhibitorzugabe vermieden werden.

15 Allerdings sind bisher keine Methoden zur gentechnischen, d. h. gezielten gentechnischen Veränderung von *Blakeslea*, insbesondere *Blakeslea trispora* bekannt.

Als Methode zur Herstellung von gentechnisch veränderten Pilzen wurde
20 in einigen Fällen die *Agrobacterium*-vermittelte Transformation erfolgreich eingesetzt. So sind z. B. folgende Organismen durch *Agrobakterien* transformiert worden: *Saccharomyces cerevisiae* (Bundock et al., 1995, EMBO Journal, 14:3206–3214), *Aspergillus awamori*, *Aspergillus nidulans*, *Aspergillus niger*, *Colletotrichum gloeosporioides*, *Fusarium solani pisi*,
25 *Neurospora crassa*, *Trichoderma reesei*, *Pleurotus ostreatus*, *Fusarium graminearum* (van der Toorren et al., 1997, EP 870835), *Agraricus bisporus*, *Fusarium venenatum* (de Groot et al., 1998, Nature Biotechnol. 16:839–842), *Mycosphaerella graminicola* (Zwiers et al. 2001, Curr. Genet. 39:388–393), *Glarea lozoyensis* (Zhang et al., 2003, Mol. Gen. Ge-

nomics 268:645–655), *Mucor miehei* (Monfort et al. 2003, FEMS Microbiology Lett. 244:101 – 106).

5 Von Interesse ist besonders eine homologe Rekombination, bei der zwischen der einzuführenden DNA und der Zell-DNA möglichst viele Sequenzhomologien bestehen, so dass eine ortsspezifische Einführung bzw. Ausschaltung von genetischer Information im Genom des Empfängerorganismus möglich ist. Andernfalls wird die Spender-DNA durch illegitime bzw. nicht-homologe Rekombination ins Genom des Empfängerorganismus integriert, was nicht ortsspezifisch erfolgt.

Eine durch *Agrobacterium* vermittelte Transformation und anschließende homologe Rekombination der transferierten DNA wurde bisher bei folgenden Organismen nachgewiesen: *Aspergillus awamori* (Gouka et al. 1999, Nature Biotech 17:598-601), *Glarea lozoyensis* (Zhang et al., 2003, Mol. Gen. Genomics 268:645-655), *Mycosphaerella graminicola* ((Zwiers et al. 2001, Curr. Genet. 39:388-393).

20 Als weitere Methode zur Transformation von Pilzen ist die Elektroporation bekannt. Die integrative Transformation von Hefe durch Elektroporation wurde von Hill, Nucl. Acids. Res. 17:8011 gezeigt. Für filamentöse Pilze wurde die Transformation durch Chakaborty und Kapoor beschrieben (1990, Nucl. Acids. Res. 18:6737).

25 Eine „biolistische“ Methode, d.h. die Übertragung von DNA durch Beschuss von Zellen mit DNA-beladenen Partikeln wurde beispielsweise für *Trichoderma harzianum* und *Gliocladium virens* beschrieben (Lorito et al. 1993, Curr. Genet. 24:349–356).

Diese Methoden konnten bisher jedoch nicht erfolgreich zur gezielten genetischen Veränderung von *Blakeslea* und insbesondere *Blakeslea trispora* eingesetzt werden.

- 5 Eine besondere Schwierigkeit bei der Herstellung von gentechnisch veränderten *Blakeslea* und *Blakeslea trispora*, ist die Tatsache, dass deren Zellen in allen Stadien des sexuellen und des vegetativen Zellzyklus mehrkernig sind. In Sporen von *Blakeslea trispora* Stamm NRRL2456 und NRRL2457 wurden z. B. im Durchschnitt 4,5 Kerne pro Spore nachgewiesen (Metha und Cerdá-Olmedo, 1995, Appl. Microbiol. Biotechnol. 42:836–
10 838). Dies hat zur Folge, dass die gentechnische Veränderung in aller Regel nur in einem oder wenigen Kernen vorliegt, die Zellen also heterokaryotisch sind.
- 15 Sollen die genetisch veränderten *Blakeslea*, insbesondere *Blakeslea trispora* zur Produktion eingesetzt werden, so ist es insbesondere bei einer Gendeletion wichtig, dass in den Produktionsstämmen die gentechnische Veränderung in allen Kernen vorliegt, so dass eine stabile und hohe Syntheseleistung ohne Nebenprodukte möglich wird. Die Stämme müssen
20 folglich in Bezug auf die gentechnische Veränderung homokaryotisch sein.

Lediglich für *Phycomyces blakesleeanus* ist ein Verfahren beschrieben worden, um homokaryotische Zellen zu erzeugen (Roncero et al., 1984, Mutat. Res. 125:195). Durch Zugabe des mutagenen Agens MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) werden nach dem dort beschriebenen
25 Verfahren Kerne in den Zellen eliminiert, so dass statistisch eine gewisse Anzahl von Zellen mit nur noch einem funktionellem Kern vorliegt. Die Zellen werden dann einer Selektion unterzogen, in der nur einkernige Zellen mit einem rezessiven Selektionsmarker zu einem Mycel auswachsen können. Die Nachkommen dieser selektierten Zellen sind mehrkernig und
30 homokaryotisch. Ein rezessiver Selektionsmarker für *Phycomyces blakes-*

leanus ist z. B. dar. dar⁺-Stämme nehmen das toxische Riboflavin-Analog 5-Carbon-5-deazariboflavin auf; dar⁻-Stämme dagegen nicht (Delbrück et al. 1979, Genetics 92:27). Die Selektion von rezessiven Mutanten erfolgt durch Zugabe von 5-Carbon-5-deazariboflavin (DARF).

5

Allerdings ist dieses Verfahren nicht für *Blakeslea*, insbesondere *Blakeslea trispora* bekannt und insbesondere nicht mit im Zusammenhang mit einer Transformation oder der Produktion von Carotinoiden oder deren Vorstufen beschrieben worden.

10

Auch die Isolierung aus natürlichen Quellen wird durchgeführt. Beispielsweise ist es für die Gewinnung von Phytoen bekannt, ein Gemisch aus Carotinoiden, Vitamin E und anderen Komponenten, welches auch Phytoen enthält, aus Tomaten, Karotten oder Palmöl usw. zu extrahieren.

15

Problematisch ist hierbei die Trennung der einzelnen Carotinoide voneinander. So ist beispielsweise das Phytoen nach diesem Verfahren nicht in reiner Form erhältlich. Insbesondere ist die natürlich vorkommende Menge der Carotinoiden in den Pflanzen gering.

20

Fermentative Verfahren sind demgegenüber technisch verhältnismäßig einfach und basieren auf kostengünstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden können dann wirtschaftlich attraktiv und wettbewerbsfähig zur chemischen Synthese sein, wenn die Produktivität der bisherigen fermentativen Verfahren gesteigert würde oder

25

neue Carotinoide auf Basis der bekannten Produktionsorganismen hergestellt werden könnten. Problematisch bei der fermentativen Herstellung von Carotinoiden sind allerdings die Aufbereitungsverfahren, die nur geringe Mengen an hochreinen Carotinoiden bereitstellen. Zudem sind dafür meist aufwendige Vielschritt-Prozesse ggf. unter Verwendung großer Lö-

30

sungsmittelmengen erforderlich. So fallen große Mengen Abfall an oder es

muss ein hoher Aufwand zur Wiederverwertung (Recycling) betrieben werden.

Die Produktion von Carotinoiden durch verschiedene Mikroorganismen ist an sich bekannt. So ist z. B. in der WO 00/13654 A2 offenbart, ein Gemisch aus Phytoen und Phytofluoren aus Algen der Art *Dunaliella* sp. zu extrahieren. Auch nach diesem Verfahren ist das Phytoen nicht in reiner Form erhältlich und muss von den anderen Produkten getrennt werden. Zudem handelt es sich um gentechnisch unveränderte Algen, deren Biosynthese mittels eines hinzugefügten Inhibitors beeinflusst werden muss.

Blakeslea trispora als Produktionsorganismus für β -Carotin ist auch aus der WO 98/03480 A1. bekannt. Hier werden β -Carotin Kristalle aus Biomasse von *Blakeslea trispora* mittels Extraktion erhalten. Allerdings müssen in dem beschriebenen Verfahren große Mengen unterschiedlicher Lösungsmittel eingesetzt werden, um Kristalle mit hoher Reinheit durch mehrere Extraktions- und Waschschrte zu erhalten. Auch sind die erhaltenen Mengen β -Carotin bezogen auf die eingesetzte Menge Biomasse klein.

Aus der WO 01/83437 A1 ist ein Verfahren zur Extraktion von Astaxanthin aus Hefe bekannt, bei dem die Kulturbrühe zur Sterilisation und zum Zellaufschluss mit Mikrowellenstrahlung behandelt wird. Der Zellaufschluss mittels Mikrowellenstrahlung ist danach nötig, um Astaxanthin aus Hefe zu gewinnen, ohne es dabei zu zerstören. Anschließend soll Astaxanthin mittels Methanol, Ethanol oder Aceton oder deren Mischungen extrahiert werden. Hierzu sind allerdings große Mengen Lösungsmittel (5 bis 20 Teile Lösungsmittel auf 1 Teil Suspension) und ein langer Zeitraum (24h) erforderlich. Zudem sind keine Reinheiten des Astaxanthins angegeben und die erhaltenen Mengen sind klein. Versuche der Anmelderin und andere Veröffentlichungen bestätigen jedoch, dass eine Extraktion mittels Methanol oder Ethanol nicht durchführbar ist.

Aus der WO 98/50574 ist ebenfalls die Isolierung von Carotinoid Kristallen aus Biomasse von Mikroorganismen bekannt, wobei hiernach im Gegensatz zur WO 01/83437 A1 Methanol, Ethanol, Aceton nur zum Entfernen
5 von Lipiden aus der Biomasse d. h. zum Waschen verwendet werden kann. Als Lösungsmittel zur Extraktion von Carotinoiden wird demnach Ethylacetat, Hexan oder ein Öl verwendet. Anschließend sind mehrere Reinigungs- und Waschschriffe mit großen Mengen Ethanol und Wasser nötig, wobei lediglich eine Reinheit von 93,3 % bei einer Ausbeute von 35
10 % erreicht wird.

Die WO 03/038064 A2 beschreibt die fermentative Produktion von Lycopin durch Co-Kultivierung von mutiertem *Blakeslea trispora* Paarungstyp (–) und *Blakeslea trispora* Paarungstyp (+), die ohne Zusatz von Inhibitoren
15 der Carotinoid Biosynthese Lycopin herstellen. Die Erzeugung der zur Fermentation eingesetzten Mutante wird durch unselektive chemische Mutation und anschließendes Screening vorgenommen. Die Aufarbeitung der Kulturbrühe erfolgt mittels Zellaufschluss und anschließender Reinigung mit unterschiedlichen wässrigen Medien mit verschiedenem Salzgehalt
20 und pH-Wert und mit Wasser nicht mischbaren organischen Lösungsmitteln wie Ethylacetat, Hexan und 1- Butanol zur Entfernung von Lipiden. Alternativ ist eine Extraktion mittels großer Mengen Ethylacetat beschrieben. Angaben zur Reinheit fehlen. Da Ethylacetat und Hexan Lösungsmittel für Lycopin sind, ist davon auszugehen, dass ein Teil des Lycopins
25 herausgewaschen und so die theoretische mögliche Ausbeute verringert wird.

Auch aus der WO 01/55100 A1 ist die Isolierung von Carotinoiden allgemein bzw. β -Carotin im speziellen aus der Biomasse durch Anwendung
30 mehrerer Wasch- und Reinigungsschritte auf die aufgeschlossene Biomasse ohne Extraktion mittels Lösungsmittel beschrieben. Hierzu wird

aufgeschlossene Biomasse von *Blakeslea trispora* mit Wasser, Lauge, Säure, Butanol und Ethanol gewaschen, so daß eine große Zahl unterschiedlicher Lösungsmittel und wässriger Medien verwendet werden muss. Die Reinheit des erhaltenen β -Carotins beträgt 96 – 98 %. Angaben zur Ausbeute fehlen jedoch.

Die WO 97/36996 A2 beschreibt allgemein ein Verfahren zur Isolierung von Substanzen (u. a. Carotinoide) aus Mikroorganismen, wobei die Substanzen aus der Biomasse mittels Fest/Flüssig-Extraktion isoliert werden. Ein Zellaufschluß soll hierbei nicht nötig sein, jedoch muss die Biomasse zunächst durch Extrusion in eine granuliert, poröse Gestalt gebracht werden. Wie nur Carotinoide isoliert werden können und wie deren Reinheit bzw. Ausbeute ist, ist nicht angegeben. Der Rückstand der Extrusion kann anschließend als Futtermittelzusatz verwendet werden.

In allen oben beschriebenen Verfahren müssen große Mengen Lösungsmittel zur Extraktion eingesetzt werden, um die isolierte Menge an Carotinoid durch vollständige Extraktion zu erhöhen, und/oder große Mengen wässriger Medien zur Reinigung und zum Waschen eingesetzt werden. Dies bedingt hohe Kosten und aufwendige Maßnahmen zur Wiederverwendung bzw. ggf. Abfälle.

Zudem werden die nahrhafte Kulturbrühe und die darin enthaltene Biomasse nach Extraktion bzw. Isolierung der Carotinoide als Abfall behandelt. Die oben angegebenen Verfahren haben neben diesen vordergründigen Nachteilen einen entscheidenden weiteren Nachteil. Es ist nämlich danach notwendig, die Carotinoide den Nahrungsmitteln nachträglich zuzusetzen, d. h. sie sind nicht Bestandteil der Nahrungsmittel an sich bzw. nicht in ausreichender Menge. Von großem Vorteil wäre daher, wenn der Gehalt an Carotinoiden in den Nahrungsmitteln bereits durch die eigentlichen Nahrungsmittel selbst gedeckt würde.

Es ist ebenfalls nötig die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin zu ermöglichen, die von den Wildtypen der Mikroorganismen bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Aufgabe der Erfindung ist es gentechnisch veränderte Zellen von Blakeslea-Stämmen, insbesondere *Blakeslea trispora* bereitzustellen, die Carotinoide oder deren Vorstufen, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin produzieren. Zudem soll das Verfahren die Steigerung der Carotinoid-Produktivität der veränderten Zellen gegenüber den korrespondierenden Wildtypen erlauben. Ferner soll das Verfahren die Erzeugung neuer Zellen oder aus ihnen bestehendes Mycel erlauben, die sich für die Verwendung zur Herstellung von Carotinoiden oder deren Vorstufen eignen, die bisher nicht in wirtschaftlich interessanten Mengen aus den natürlich vorkommenden Pilzen gewinnbar waren, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin. Das Verfahren soll dabei eine gentechnische Veränderung von *Blakeslea*-Stämmen, insbesondere *Blakeslea trispora* möglich machen und die Herstellung homokaryotischer gentechnisch veränderter Produktions-Stämme erlauben.

Des weiteren soll das Verfahren die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle, insbesondere Astaxanthin oder Zeaxanthin und Phytoen oder Bixin ermöglichen, die von den Wildtypen der Mikroorganismen bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Ferner ist es Aufgabe der vorliegenden Erfindung ein Verfahren zur Herstellung von Carotinoiden aus gentechnisch veränderte Zellen von Blakeslea-Stämmen, insbesondere Blakeslea trispora, zur Verfügung zu stellen, welches den Einsatz geringerer Lösungsmittelmengen erlaubt und im wesentlichen ohne Abfälle auskommt und zudem eine hohe Reinheit und höhere Ausbeuten erlaubt.

In diesem Zusammenhang soll ein möglichst großer Anteil der im Fermenter vorliegenden Nährstoffe, sowohl Carotinoide als auch weitere sich in den Mikroorganismen befindende, verwertet werden.

Somit ist es auch Aufgabe der vorliegenden Erfindung ein Verfahren zur Herstellung eines Carotinoid-haltigen Nahrungsmittels bereitzustellen, wobei das Nahrungsmittel selbst den Bedarf an Carotinoiden ohne Zusätze deckt. Insbesondere soll der Nährstoffgehalt der nach dem Verfahren erhältlichen Nahrungsmittel gegenüber den bisher erhältlichen Nahrungsmitteln zumindest gleichwertig sein. Ferner soll das Verfahren die effiziente Verwertung der produzierten Carotinoide ermöglichen.

Diese Aufgabe wird durch ein Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderten Organismen der Gattung Blakeslea gelöst, umfassend

- (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und
- (iii) Selektion und Vermehrung der gentechnisch veränderten Zelle oder Zellen,
- (iv) Kultivierung der gentechnisch veränderten Zellen,

- (v) Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe.

- 5 Mit der erfindungsgemäßen Methode ist es möglich, *Blakeslea* gezielt und stabil genetisch zu verändern, um so Mycel aus Zellen mit einheitlichen Kernen zu gewinnen, das Carotinoide oder deren Vorstufen, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin produziert. Vorzugsweise handelt es sich um Zellen von
- 10 Pilzen der Art *Blakeslea trispora*. Die produzierten Carotinoiden oder deren Vorstufen sind dabei im wesentlichen frei von Verunreinigungen erhältlich und es können hohe Konzentrationen der Carotinoiden oder deren Vorstufen im Kulturmedium erzielt werden.
- 15 Unter Transformation wird die Übertragung einer genetischen Information in den Organismus, insbesondere Pilz verstanden. Darunter sollen alle dem Fachmann bekannten Möglichkeiten zur Einschleusung der Information, insbesondere DNA fallen, z. B. Beschuss mit DNA-beladenen Partikeln, Transformation mittels Protoplasten, Mikroinjektion von DNA, Elektroporation, Konjugation oder Transformation kompetenter Zellen, Chemikalien oder Agrobakterien vermittelte Transformation. Als genetische In-
- 20 formation werden ein Genabschnitt, ein Gen oder mehrere Gene verstanden. Die genetische Information kann z. B. mit Hilfe eines Vectors oder als freie Nukleinsäure (z. B. DNA, RNA) und auf sonstige Weise in die Zellen
- 25 eingebracht und entweder durch Rekombination ins Wirtsgenom eingebaut oder in freier Form in der Zelle vorliegen. Besonders bevorzugt ist hierbei die homologe Rekombination.

Bevorzugte Transformationsmethode ist die *Agrobacterium tumefaciens*-

30 vermittelte Transformation. Hierzu wird zunächst die zu transferierende

Spender-DNA in einen Vektor eingefügt, der (i) flankierend zu der zu transferierenden DNA die T-DNA-Enden trägt, der (ii) einen Selektionsmarker enthält und der (iii) ggf. Promotoren und Terminatoren für die Genexpression der Spender-DNA aufweist. Dieser Vektor wird in einen Agrobacterium-tumefaciens-Stamm übertragen, der ein Ti-Plasmid mit den vir-
5 Genen enthält. vir-Gene sind für den DNA-Transfer in Blakeslea verantwortlich. Mit diesem Zwei-Vektor-System wird die DNA von Agrobacterium in Blakeslea übertragen. Hierzu werden die Agrobakterien zunächst in Gegenwart von Acetosyringone inkubiert. Acetosyringone induziert die vir-
10 Gene. Anschließend werden Sporen von Blakeslea trispora zusammen mit den induzierten Zellen von Agrobacterium tumefaciens auf Acetosyringone-haltigem Medium inkubiert und dann auf Medium übertragen, das eine Selektion der Transformanten, d.h. der gentechnisch veränderten Stämme von Blakeslea ermöglicht.

15

Der Begriff Vector wird in der vorliegenden Anmeldung als eine Bezeichnung für ein DNA-Molekül verwendet, das zum Einschleusen und ggf. zur Vermehrung von Fremd-DNA in eine Zelle dient (siehe auch "Vector" in Römpp Lexikon Chemie – CDROM Version 2.0, Stuttgart/New York: Georg Thieme Verlag 1999). In der vorliegenden Anmeldung sollen unter
20 dem begriff "Vector" auch Plasmide, Cosmide usw. verstanden werden, die dem gleichen Zweck dienen.

Unter Expression wird in der vorliegenden Anmeldung die Übertragung
25 einer genetischen Information ausgehend von DNA oder RNA in ein Gen-Produkt (hier vorzugsweise Enzyme zur Herstellung von Carotinoiden und insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin) verstanden und soll auch den Begriff der Überexpression beinhalten, womit eine verstärkte Expression gemeint ist,
30 so dass ein bereits in der nicht transformierten Zelle (Wildtyp) hergestell-

tes Genprodukt verstärkt produziert wird oder einen großen Teil des gesamten Gehaltes der Zelle ausmacht.

Unter gentechnische Veränderung soll die Einschleusung genetischer Information in einen Empfängerorganismus, so dass diese stabil exprimiert und bei der Zellteilung weitergegeben wird, verstanden werden. In diesem Zusammenhang ist die Homokaryotisierung, die Herstellung von Zellen, die nur einheitliche Kerne enthalten, d. h. Kerne mit gleichem genetischem Informationsgehalt.

Diese Homokaryotisierung ist nur notwendig, wenn die durch Transformation eingeführte genetische Information rezessiv vorliegt, d. h. nicht zur Ausprägung gelangt. Führt die Transformation aber zu einem dominanten Vorliegen der genetischen Information, d. h. wird sie ausgeprägt, so ist eine Homokaryotisierung nicht unbedingt nötig.

Vorzugsweise wird zur Homokaryotisierung eine Selektion der einkernigen Sporen durchgeführt. Von Natur aus ist ein geringer Anteil der Sporen von *Blakeslea trispora* einkernig, so dass sich diese ggf. nach spezifischer Markierung z. B. Färbung der Zellkerne aussortieren lassen. Dies wird bevorzugterweise mittels FACS (Fluorescence Activated Cell Sorting) anhand der geringeren Fluoreszenz der einkernigen Zellen durchgeführt.

Alternativ kann zur Homokaryotisierung zunächst eine Kernreduktion durchgeführt werden. Hierzu kann ein mutagenes Agens eingesetzt werden, wobei es sich insbesondere um N-Methyl-N'-nitro-nitrosoguanidin (MNNG) handelt. Auch die Verwendung von energiereichen Strahlen, wie UV- oder Röntgen-Strahlen zur Kernreduktion ist möglich. Anschließend kann zur Selektion auf das FACS Verfahren oder rezessive Selektionsmarker zurückgegriffen werden.

Unter Selektion wird die Auswahl von Zellen verstanden, deren Kerne dieselbe genetische Information beinhalten, d. h. Zellen die die gleichen Eigenschaften aufweisen, wie Resistenzen oder die Herstellung bzw. vermehrte Herstellung eines Produktes. In der Selektion werden neben der
5 FACS Methode bevorzugt 5-Carbon-5-deazariboflavin (DARF) und Hygromycin (hyg) oder 5'-Fluororotat (FOA) und Uracil eingesetzt.

Der in der Transformation (i) eingesetzte Vector kann derart gestaltet sein, dass die im Vector enthaltene genetische Information in das Genom mindestens einer Zelle integriert wird. Dabei kann genetische Information in
10 der Zelle ausgeschaltet werden. Dies kann direkt, d. h. durch eine Deletion erfolgen. Es ist aber auch möglich, daß der in der Transformation (i) eingesetzte Vector derart ausgestaltet ist, dass die im Vector enthaltene genetische Information in der Zelle exprimiert wird, d. h. genetische Information
15 eingefügt wird, die im korrespondierenden Wildtyp nicht vorhanden ist oder die durch die Transformation verstärkt bzw. überexprimiert wird und deren Produkt das Gen ausschaltet. Die eingeführte genetische Information kann aber auch indirekt eine genetische Information in der Zelle ausschalten, z. B. durch Produktion eines Inhibitors.

20

Der eingesetzte Vector enthält genetische Informationen oder Teile der genetischen Information zur Herstellung von Carotinoiden oder deren Vorstufen, insbesondere Carotinen oder Xanthophyllen oder deren Vorstufen. Der eingesetzte Vector enthält vorzugsweise genetische Informationen zur
25 Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, β -Carotin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, Lutein, Bixin oder Phytoen. Ganz besonders bevorzugt enthält der Vector Informationen zur Herstellung von Bixin, Phytoen, Canthaxanthin, Astaxanthin oder Zeaxanthin.

30

Der Vector kann beliebige genetische Informationen zur genetischen Veränderungen von Organismen der Gattung *Blakeslea* enthalten.

5 Unter „genetischer Information“ werden vorzugsweise Nukleinsäuren verstanden, deren Einbringung in den Organismus der Gattung *Blakeslea* zu einer genetischen Veränderung in Organismen der Gattung *Blakeslea*, also beispielsweise zu einer Verursachung, Erhöhung oder Reduzierung von Enzymaktivitäten im Vergleich zum Ausgangsorganismus führen.

10 Der Vector kann beispielsweise genetische Information zur Herstellung lipophiler Substanzen enthalten wie z.B. Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder genetische Information zur Herstellung hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und
15 wasserlösliche Vitaminen, Provitamine und Cofaktoren.

Bevorzugterweise enthält der eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder Xanthophyllen oder deren Vorstufen.

20 Bevorzugterweise enthält der Vektor genetische Information, die eine Lokalisierung der Carotinoidbiosynthese-Enzyme in dem Zellkompartiment bewirkt, in dem die Carotinoidbiosynthese stattfindet.

25 Besonders bevorzugt sind genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Anthonixanthin, Adonirubin, Canthaxanthin, 3- und 3'-Hydroxyechinenon, Lycopin, Lutein, β -Carotin, Phytoen und/oder Phytoflu. Ganz besonders bevorzugt sind genetische Informationen zur Herstellung von Phytoen, Bixin, Lycopin,
30 Zeaxanthin, Canthaxanthin und/oder Astaxanthin.

Entsprechend werden in einer bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die über eine erhöhte Syntheserate für Zwischenprodukte der Carotinoidbiosynthese verfügen und folglich eine erhöhte Produktivität für Endprodukte der Carotinoidbiosynthese aufweisen. Zur Erhöhung der Syntheserate für Zwischenprodukte der Carotinoidbiosynthese werden insbesondere die Aktivitäten der Enzyme 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase (HMG-CoA-Reduktase), Isopentenylpyrophosphat-Isomerase und Geranylpyrophosphatsynthase gesteigert.

10

Entsprechend werden in einer besonders bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die gegenüber dem Wildtyp eine erhöhte HMG-CoA-Reduktase-Aktivität aufweisen.

15 Unter HMG-CoA-Reduktase-Aktivität wird die Enzymaktivität einer HMG-CoA-Reduktase (3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase) verstanden.

Unter einer HMG-CoA-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A in Mevalonat umzuwandeln.

20

Dementsprechend wird unter HMG-CoA-Reduktase-Aktivität die in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. gebildete Menge Mevalonat verstanden.

25

Bei einer erhöhten HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase die umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. die gebildete Menge Mevalonat erhöht.

30

Vorzugsweise beträgt diese Erhöhung der HMG-CoA-Reduktase-Aktivität mindestens 5%, weiter bevorzugt mindestens 20%, weiter bevorzugt mindestens 50%, weiter bevorzugt mindestens 100%, besonders bevorzugt mindestens 300%, noch bevorzugter mindestens 500%, insbesondere
5 mindestens 600% der HMG-CoA-Reduktase-Aktivität des Wildtyps.

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp durch eine Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase.

10

In einer besonders bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens erfolgt die Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase indem man ein Nukleinsäurekonstrukt, enthaltend eine Nukleinsäure codierend eine HMG-CoA-Reduktase
15 in den Organismus einbringt, deren Expression in dem Organismus, verglichen mit dem Wildtyp, einer reduzierten Regulation unterliegt.

Unter einer reduzierten Regulation verglichen mit dem Wildtyp, wird eine im Vergleich zum vorstehend definierten Wildtyp verringerte, vorzugsweise
20 keine Regulation auf Expressions- oder Proteinebene verstanden.

Die reduzierte Regulation kann vorzugsweise durch einen im Nukleinsäurekonstrukt mit der kodierenden Sequenz funktionell verknüpften Promotor erreicht werden, der in dem Organismus verglichen mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.
25

Beispielsweise unterliegen die Promotoren ptef1 aus *Blakeslea trispora* und pgpdA aus *Aspergillus nidulans* nur einer reduzierten Regulation und sind daher insbesondere als Promotoren bevorzugt.

30

Diese Promotoren zeigen eine annähernd konstitutive Expression in *Blakeslea trispora*, so dass die transkriptionelle Regulation nicht mehr über die Intermediate der Carotinoidbiosynthese abläuft.

- 5 Die reduzierte Regulation kann in einer weiteren bevorzugten Ausführungsform dadurch erreicht werden, dass man als Nukleinsäure codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt.

10

- Besonders bevorzugt ist die Verwendung einer Nukleinsäure, die nur den katalytischen Bereich der HMG-CoA-Reduktase kodiert (trunkierte (t-)HMG-CoA-Reduktase). Die für die Regulation verantwortliche Membran-Domäne fehlt. Die verwendete Nukleinsäure unterliegt somit einer reduzierten Regulation und führt zu einer Erhöhung der Genexpression der HMG-CoA-Reduktase.

15

- In einer besonders bevorzugten Ausführungsform bringt man Nukleinsäuren in *Blakeslea trispora* ein, welche die Sequenz SEQ ID. NO. 75 enthalten.

20

- Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Homologievergleiche der Sequenzen aus Datenbanken mit der SEQ ID. NO. 75 leicht finden.

25

- Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich weiterhin beispielsweise ausgehend von

30

der Sequenz SEQ ID. NO. 75 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

- 5 In einer besonders bevorzugten Ausführungsform wird die reduzierte Regulation dadurch erreicht, dass man als Nukleinsäure codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt und einen Promotor
10 verwendet, der in dem Organismus, verglichen mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.

Entsprechend wird in einer bevorzugten Variante der Erfindung durch die Transformation die Genexpression der Phytoendesaturase ausgeschaltet,
15 so dass das von den Organismen produzierte Phytoen gewonnen werden kann. Der in der Transformation (i) eingesetzte Vector umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Phytoendesaturase, insbesondere carB aus *Blakeslea trispora* mit der SEQ ID NO: 69.

20

Entsprechend wird in einer bevorzugten Variante der Erfindung durch Transformation die Genexpression der Lycopincyclase ausgeschaltet, so dass das von den Organismen produzierte Lycopin gewonnen werden kann. Der in der Transformation eingesetzte Vektor umfasst daher in einer
25 Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Lycopincyclase, insbesondere carR aus *Blakeslea trispora*.

- 30 In einer bevorzugten Ausführungsform werden die Organismen der Gattung *Blakeslea* beispielsweise dadurch in die Lage versetzt Xanthophylle,

wie beispielsweise Canthaxanthin, Zeaxanthin oder Astaxanthin herzustellen, Bixin oder Phytoen, indem in den genetisch veränderten Organismen der Gattung *Blakeslea* im Vergleich zum Wildtyp eine Hydroxylase-Aktivität und/oder Ketolase-Aktivität verursacht wird.

5

Der in der Transformation (i) eingesetzte Vector enthält also in einer weiteren, bevorzugten Variante der Erfindung genetische Informationen, die nach Expression eine Ketolase- und/oder Hydroxylase-Aktivität entfalten, so dass die Organismen Zeaxanthin oder Astaxanthin produzieren.

10

Unter Ketolase-Aktivität wird die Enzymaktivität einer Ketolase verstanden.

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

15

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Canthaxanthin umzuwandeln.

20

Dementsprechend wird unter Ketolase-Aktivität die in einer bestimmten Zeit durch das Protein Ketolase umgesetzte Menge β -Carotin bzw. gebildete Menge Canthaxanthin verstanden.

Unter dem Begriff "Wildtyp" wird erfindungsgemäß der entsprechende nicht genetisch veränderte Ausgangsorganismus der Gattung *Blakeslea* verstanden.

25

Je nach Zusammenhang kann unter dem Begriff "Organismus" der Ausgangsorganismus (Wildtyp) der Gattung *Blakeslea* oder ein erfindungs-

30

gemäß, genetisch veränderter Organismus der Gattung *Blakesleaa* oder beides verstanden werden.

Vorzugsweise wird unter "Wildtyp" für die Verursachung der Ketolase-Aktivität und für die Verursachung der Hydroxylase-Aktivität jeweils eine Referenz Organismus verstanden.

Dieser Referenzorganismus der Gattung *Blakeslea* ist *Blakeslea trispora* ATCC 14271 oder ATCC 14272, die sich lediglich im Paarungstyp unterscheiden.

Die Bestimmung der Ketolase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen der Gattung *Blakesleaa* und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Bestimmung der Ketolase-Aktivität in Organismen der Gattung *Blakeslea* erfolgt in Anlehnung an die Methode von Frazer et al., (J. Biol. Chem. 272(10): 6128-6135, 1997). Die Ketolase-Aktivität in Extrakten wird mit den Substraten beta-Carotin und Canthaxanthin in Gegenwart von Lipid (Sojalecithin) und Detergens (Natriumcholat) bestimmt. Substrat/Produkt-Verhältnisse aus den Ketolase-Assays werden mittels HPLC ermittelt.

Der erfindungsgemäße genetisch veränderte Organismus der Gattung *Blakesleaa* weist in dieser, bevorzugten Ausführungsform im Vergleich zum genetisch nicht veränderten Wildtyp eine Ketolase-Aktivität auf und ist somit vorzugsweise in der Lage, transgen eine Ketolase zu exprimieren.

In einer weiter bevorzugten Ausführungsform erfolgt die Verursachung der Ketolase-Aktivität in den Organismen der Gattung *Blakesleaa* durch Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase.

In dieser bevorzugten Ausführungsform erfolgt die Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase vorzugsweise durch Einbringen von Nukleinsäuren, die Ketolasen kodieren in die Ausgangsorganismus der Gattung *Blakesleaa*.

Dazu kann prinzipiell jedes Ketolase-Gen, also jede Nukleinsäuren die eine Ketolase codiert verwendet werden.

Alle in der Beschreibung erwähnten Nukleinsäuren können beispielsweise eine RNA-, DNA- oder cDNA-Sequenz sein.

Bei genomischen Ketolase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall das der Wirtsorganismus der Gattung *Blakesleaa* nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechenden Ketolase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

Beispiele für Nukleinsäuren, kodierend eine Ketolase und die entsprechenden Ketolasen, die im erfindungsgemäßen Verfahren verwendet werden können sind beispielsweise Sequenzen aus:

Haematoccus pluvialis, insbesondere aus *Haematoccus pluvialis* Flotow em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 11, Protein SEQ ID NO: 12),

Haematoccus pluvialis, NIES-144 (Accession NO: D45881; Nukleinsäure: SEQ ID NO: 13, Protein SEQ ID NO: 14),

Agrobacterium aurantiacum (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),

5 *Aliccaligenes spec.* (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 17, Protein SEQ ID NO: 18),

Paracoccus marcusii (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20).

10 *Synechocystis sp.* Strain PC6803 (Accession NO: NP442491; Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22).

Bradyrhizobium sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24).

15 *Nostoc sp.* Strain PCC7120 (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26),

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No. NZ_AABC01000195, Basenpaar 55,604 bis 55,392 (SEQ ID NO: 27); Protein: Acc.-No. ZP_00111258 (SEQ ID NO: 28) (als putatives Protein annotiert),

20 *Nostoc punctiforme* ATTC 29133, Nukleinsäure: Acc.-No. NZ_AABC01000196, Basenpaar 140,571 bis 139,810 (SEQ ID NO: 29), Protein: (SEQ ID NO: 30) (nicht annotiert),

30 Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene, die im erfindungsgemäßen Verfahren verwendet werden können, lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Aminosäuresequenzen oder der

entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen und insbesondere mit den Sequenzen SEQ ID NO: 12, 26 und/oder 33 leicht auffinden.

- 5 Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 12, 26 und/oder 30 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich be-
- 10 kannter Weise leicht auffinden.

Die Hybridisierung kann unter moderaten (geringe Stringenz) oder vorzugsweise unter stringenten (hohe Stringenz) Bedingungen erfolgen.

- 15 Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben.

- 20 Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit 2X SSC bei 50°C) und solchen mit hoher Stringenz (mit 0.2X SSC bei 50°C, bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M Natriumchlorid, pH 7.0).
- 25

Darüberhinaus kann die Temperatur während des Waschschrilles von moderaten Bedingungen bei Raumtemperatur, 22°C, bis zu stringenten Bedingungen bei 65°C angehoben werden.

Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50% Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt.

Einige beispielhafte Bedingungen für Hybridisierung und Waschschrift sind infolge gegeben:

10

- (1) Hybridisierungsbedingungen mit zum Beispiel
 - (i) 4X SSC bei 65°C, oder
 - (ii) 6X SSC bei 45°C, oder
 - (iii) 6X SSC bei 68°C, 100 mg/ml denaturierter Fischsperma-DNA, oder
 - 15 (iv) 6X SSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachs-sperma-DNA bei 68°C, oder
 - (v) 6XSSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachs-sperma-DNA, 50 % Formamid bei 42°C, oder
 - (vi) 50 % Formamid, 4X SSC bei 42°C, oder
 - 20 (vii) 50 % (vol/vol) Formamid, 0.1 % Rinderserumalbumin, 0.1 % Ficoll, 0.1 % Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6.5, 750 mM NaCl, 75 mM Natriumcitrat bei 42°C, oder
 - (viii) 2X oder 4X SSC bei 50°C (moderate Bedingungen), oder
 - (ix) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (moderate Be-
25 dingungen).

- (2) Waschschriffe für jeweils 10 Minuten mit zum Beispiel
 - (i) 0.015 M NaCl/0.0015 M Natriumcitrat/0.1 % SDS bei 50°C, oder
 - (ii) 0.1X SSC bei 65°C, oder
 - 30 (iii) 0.1X SSC, 0.5 % SDS bei 68°C, oder
 - (iv) 0.1X SSC, 0.5 % SDS, 50 % Formamid bei 42°C, oder

- (v) 0.2X SSC, 0.1 % SDS bei 42°C, oder
- (vi) 2X SSC bei 65°C (moderate Bedingungen).

In einer bevorzugten Ausführungsform der erfindungsgemäßen genetisch
5 veränderten Organismen der Gattung *Blakeslea* bringt man Nukleinsäuren
ein, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID
NO: 12 oder eine von dieser Sequenz durch Substitution, Insertion oder
Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von
mindestens 20 %, vorzugsweise mindestens 30%, 40%, 50%, 60%, be-
vorzugt mindestens 70%, 80%, besonders bevorzugt mindestens 90%,
10 insbesondere 91%, 92%; 93%, 94%, 95%, 96%, 97%, 98% oder 99% auf
Aminosäureebene mit der Sequenz SEQ ID NO: 12 und die enzymatische
Eigenschaft einer Ketolase aufweist.

15 Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie
vorstehend beschrieben durch Identitätsvergleich der Sequenzen aus an-
deren Organismen gefunden werden kann oder um eine künstliche Keto-
lase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 12 durch
künstliche Variation, beispielsweise durch Substitution, Insertion oder De-
20 letion von Aminosäuren abgewandelt wurde.

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen
Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthal-
tend die Aminosäuresequenz SEQ ID NO: 26 oder eine von dieser Se-
25 quenz durch Substitution, Insertion oder Deletion von Aminosäuren abge-
leitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise
mindestens 30%, 40%, 50%, 60%, bevorzugt mindestens 70%, 80%, be-
sonders bevorzugt mindestens 90%, insbesondere 91%, 92%; 93%, 94%,
95%, 96%, 97%, 98% oder 99% auf Aminosäureebene mit der Sequenz
30 SEQ ID NO: 26 und die enzymatische Eigenschaft einer Ketolase auf-
weist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 26 durch
5 künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen
10 Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 30 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise
15 mindestens 30 %, 40 %, 50 %, bevorzugt mindestens 60 %, 70 %, bevorzugter mindestens 80 %, 85 % besonders bevorzugt mindestens 90 %, insbesondere 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % auf Aminosäureebene mit der Sequenz SEQ ID NO 30 und die enzymatische Eigenschaft einer Ketolase aufweist.

20 Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 30 durch
25 künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

Unter dem Begriff "Substitution" ist in der Beschreibung der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren zu verstehen. Bevorzugt werden sog. konservative Austausche durchge-
30 führt, bei denen die ersetzte Aminosäure eine ähnliche Eigenschaft hat

wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

Deletion ist das Ersetzen einer Aminosäure durch eine direkte Bindung.

5 Bevorzugte Positionen für Deletionen sind die Termini des Polypeptides und die Verknüpfungen zwischen den einzelnen Proteindomänen.

Insertionen sind Einfügungen von Aminosäuren in die Polypeptidkette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren

10 ersetzt wird.

Unter Identität zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinlänge verstanden, insbesondere die Identität die durch Vergleich mit Hilfe der Lasergene Software der Firma
15 DNASTAR, inc. Madison, Wisconsin (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

20 Multiple alignment parameter:

Gap penalty 10

Gap length penalty 10

Pairwise alignment parameter:

K-tuple 1

25 Gap penalty 3

Window 5

Diagonals saved 5

Unter einem Protein, das eine Identität von mindestens 20% auf Aminosäureebene mit der Sequenz SEQ ID NO: 12 oder 26 oder 30 aufweist,
30 wird dementsprechend ein Protein verstanden, das bei einem Vergleich

seiner Sequenz mit der Sequenz SEQ ID NO: 12 oder 26 oder 30, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 20 %, bevorzugt 30%, 40%, 50%, besonders bevorzugt 60%, 70%, 80%, insbesondere 85%, 90, 95% aufweist.

5

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

10

Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Blakesleaspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene von Organismen der Gattung Blakesleaa leicht ermitteln.

15

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 11 in die Organismus der Gattung ein.

20

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 25 in die Organismus der Gattung ein.

25

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 29 in die Organismus der Gattung ein.

30

Alle vorstehend erwähnten Ketolase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage,

Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Der in der Transformation (i) eingesetzte Vector umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für eine Ketolase, insbesondere der Ketolase *Nostoc punctiforme* aus mit der SEQ ID NO: 72.

Unter Hydroxylase-Aktivität die Enzymaktivität einer Hydroxylase verstanden.

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Zeaxanthin oder Cantaxanthin in Astaxanthin umzuwandeln.

Dementsprechend wird unter Hydroxylase-Aktivität die in einer bestimmten Zeit durch das Protein Hydroxylase umgesetzte Menge β -Carotin oder Cantaxanthin bzw. gebildete Menge Zeaxanthin oder Astaxanthin verstanden.

Bei einer erhöhten Hydroxylase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein

Hydroxylase die umgesetzte Menge β -Carotin oder Cantaxantin bzw. die gebildete Menge Zeaxanthin oder Astaxanthin erhöht.

Vorzugsweise beträgt diese Erhöhung der Hydroxylase-Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, bevorzugter mindestens 300 %, noch bevorzugter mindestens 500 %, insbesondere mindestens 600 % der Hydroxylase-Aktivität des Wildtyps.

Die Bestimmung der Hydroxylase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Aktivität der Hydroxylase wird nach Bouvier et al. (Biochim. Biophys. Acta 1391 (1998), 320-328) *in vitro* bestimmt. Es wird zu einer bestimmten Menge an Organismenextrakt Ferredoxin, Ferredoxin-NADP Oxidoreduktase, Katalase, NADPH sowie beta-Carotin mit Mono- und Digalaktosylglyzeriden zugegeben.

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, Keller, d'Harlingue und Camara (Xanthophyll biosynthesis: molecular and functional characterization of carotenoid hydroxylases from pepper fruits (*Capsicum annuum* L.; Biochim. Biophys. Acta 1391 (1998), 320-328):

25

Der *in-vitro* Assay wird in einem Volumen von 0.250 ml Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), 0.025 mg Ferredoxin von Spinat, 0.5 Einheiten Ferredoxin-NADP⁺ Oxidoreduktase von Spinat, 0.25 mM NADPH, 0.010 mg beta-Carotin (in 0.1 mg Tween 80 emulgiert), 0.05 mM einer Mischung von Mono- und Digalaktosylglyzeriden (1:1), 1 Einheit Katalase, 200 Mono- und Digalaktosylglyzeriden, (1:1),

30

0.2 mg Rinderserumalbumin und Organismenextrakt in unterschiedlichem Volumen. Die Reaktionsmischung wird 2 Stunden bei 30°C inkubiert. Die Reaktionsprodukte werden mit organischem Lösungsmittel wie Aceton oder Chloroform/Methanol (2:1) extrahiert und mittels HPLC bestimmt.

5

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, d'Harlingue und Camara (Molecular Analysis of carotenoid cycle inhibition; Arch. Biochem. Biophys. 346(1) (1997) 53-64):

10

Der in-vitro Assay wird in einem Volumen von 250 µl Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), unterschiedliche Mengen an Organismenextrakt, 20 nM Lycopin, 250 µg an chromoplastidärem Stromaprotein aus Paprika, 0.2 mM NADP⁺, 0.2 mM NADPH und 1 mM ATP. NADP/NADPH und ATP werden in 10 ml Ethanol mit 1 mg Tween 80 unmittelbar vor der Zugabe zum Inkubationsmedium gelöst. Nach einer Reaktionszeit von 60 Minuten bei 30°C wird die Reaktion durch Zugabe von Chloroform/Methanol (2:1) beendet. Die in Chloroform extrahierten Reaktionsprodukte werden mittels HPLC analysiert.

20

Ein alternativer Assay mit radioaktivem Substrat ist beschrieben in Fraser und Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15).

25

Die Erhöhung der Hydroxylase-Aktivität kann durch verschiedene Wege erfolgen, beispielsweise durch Ausschalten von hemmenden Regulationsmechanismen auf Expressions- und Proteinebene oder durch Erhöhung der Genexpression von Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp.

30

Die Erhöhung der Genexpression der Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp kann ebenfalls durch verschiedene

Wege erfolgen, beispielsweise durch Induzierung des Hydroxylase-Gens durch Aktivatoren oder durch Einbringen von einer oder mehreren Hydroxylase-Genkopien, also durch Einbringen mindestens einer Nukleinsäure kodierend eine Hydroxylase in den Organismus der Gattung *Blakesleaa*.

5

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der Genexpression einer Nukleinsäure kodierend eine Hydroxylase durch Einbringen von mindestens einer Nukleinsäure kodierend eine Hydroxylase in den Organismus der Gattung *Blakesleaa*.

10

Dazu kann prinzipiell jedes Hydroxylase-Gen, also jede Nukleinsäure, die eine Hydroxylase und jede Nukleinsäure, die eine β -Cyclase codiert, verwendet werden.

15 Bei genomischen Hydroxylase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall das der Wirtsorganismus nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechende Hydroxylase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

20

Ein Beispiel für ein Hydroxylase-Gen ist eine Nukleinsäure, kodierend eine Hydroxylase aus *Haematococcus pluvialis* mit der Accession No. AX038729 (WO 0061764; Nukleinsäure: SEQ ID NO: 31, Protein: SEQ ID NO: 32), aus *Erwinia uredovora* 20D3 (ATCC 19321, Accession No. D90087; Nukleinsäure: SEQ ID NO: 33, Protein: SEQ ID NO: 34) oder Hydroxylase aus *Thermus thermophilus* (DE 102 34 126.5) kodiert durch die Sequenz mit der SEQ ID NO 76.

25

sowie Hydroxylasen der folgenden Accession Nummern:
30 |emb|CAB55626.1, CAA70427.1, CAA70888.1, CAB55625.1, AF499108_1, AF315289_1, AF296158_1, AAC49443.1, NP_194300.1,

NP_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1,
AAL80006.1, AF162276_1, AAO53295.1, AAN85601.1, CRTZ_ERWHE,
CRTZ_PANAN, BAB79605.1, CRTZ_ALCSP, CRTZ_AGRAU,
CAB56060.1, ZP_00094836.1, AAC44852.1, BAC77670.1, NP_745389.1,
5 NP_344225.1, NP_849490.1, ZP_00087019.1, NP_503072.1,
NP_852012.1, NP_115929.1, ZP_00013255.1

In den erfindungsgemäßen bevorzugten transgenen Organismen der Gat-
tung *Blakeslea* liegt also in dieser bevorzugten Ausführungsform gegen-
10 über dem Wildtyp mindestens ein weiteres Hydroxylase-Gen vor.

In dieser bevorzugten Ausführungsform weist die genetisch veränderte
Organismus beispielsweise mindestens eine exogene Nukleinsäure, ko-
dierend eine Hydroxylase oder mindestens zwei endogene Nukleinsäuren,
15 kodierend eine Hydroxylase auf.

Bevorzugt verwendet man in vorstehend beschriebener bevorzugter Aus-
führungsform als Hydroxylase-Gene Nukleinsäuren, die Proteine kodieren,
enthaltend die Aminosäuresequenz SEQ ID NO: 32, 34 oder kodiert durch
20 die Sequenz mit der SEQ ID NO 76 oder eine von dieser Sequenz durch
Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Se-
quenz, die eine Identität von mindestens 30 %, vorzugsweise mindestens
50 %, bevorzugter mindestens 70%, noch bevorzugter mindestens 80 %,
am bevorzugtesten mindestens 90%, insbesondere 91%, 92%, 93%, 94%,
25 95%, 96%, 97%, 98%, 99% auf Aminosäureebene mit der Sequenz
SEQ. ID. NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO
76 und die die enzymatische Eigenschaft einer Hydroxylase aufweisen.

Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich
30 beispielsweise aus verschiedenen Organismen, deren genomische Se-
quenz bekannt ist, wie vorstehend beschrieben, durch Homologieverglei-

che der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der SEQ ID. NO: 31, 33 oder 76 leicht auffinden.

- 5 Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich weiterhin beispielsweise ausgehend von der Sequenz SEQ ID NO: 31, 33 oder 76 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, wie vorstehend beschrieben, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

10

In einer weiter besonders bevorzugten Ausführungsform werden zur Erhöhung der Hydroxylase-Aktivität Nukleinsäuren in Organismen eingebracht, die Proteine kodieren, enthaltend die Aminosäuresequenz der Hydroxylase der Sequenz SEQ ID NO: 32, 34 oder kodiert durch die Sequenz mit

15

der SEQ ID NO 76.

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

- 20 Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Organismenspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

- 25 In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ. ID. NO: 31, 33 oder 76 in den Organismus ein.

- 30 Alle vorstehend erwähnten Hydroxylase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender,

komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

10

Der in der Transformation (i) eingesetzte Vector umfasst daher in weiteren Ausführungsformen der Erfindung bevorzugterweise eine Sequenz codierend für eine Hydroxylase, insbesondere eine Hydroxylase aus *Haemotococcus pluvialis* mit der SEQ ID NO: 70 oder eine Hydroxylase aus *Erwinia uredova* mit der SEQ ID NO: 71. oder eine Hydroxylase aus *Thermus thermophilus* kodiert durch die Sequenz mit der SEQ ID NO 76.

15

Vorzugsweise wird durch die Transformation das Gen der Phytoendesaturase ausgeschaltet.

20

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise ferner die Expression regelnde und unterstützende Bereiche, insbesondere Promotoren und Terminatoren.

25

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise den *gpd* und/oder den *ptef1* Promotor und/oder den *trpC* Terminator. Diese haben sich zur Transformation der *Blakeslea* besonders bewährt. Auch der Einsatz von dem Fachmann geläufigen "inverted repeats" (IR, Römpp Lexikon der Biotechnologie 1992, Thieme Verlag Stuttgart, Seite 407 "Invers repetitive Sequenzen") zur Regelung der Expression bzw. Transkription liegt im Rahmen der Erfindung.

30

Vorteilhafterweise weist der im Vector eingesetzte gpd Promotor die Sequenz SEQ ID NO: 1 auf. Vorteilhafterweise weist der im Vector eingesetzte trpC Terminator die Sequenz SEQ ID NO: 2 auf.
5 Vorteilhafterweise weist der im Vector eingesetzte ptef1 Promotor die Sequenz SEQ ID NO: 35 auf.

Insbesondere werden dabei der gpd Promotor und der trpC Terminator aus *Aspergillus nidulans* und der ptef1 Promotor aus *Blakeslea trispora*
10 eingesetzt.

Insbesondere enthält der in der Transformation (i) eingesetzte Vector ein Resistenzgen. Bevorzugterweise handelt es sich um ein Hygromycin-Resistenzgen (hph), insbesondere eines aus *E. coli*. Dieses Resistenzgen
15 hat sich bei dem Nachweis der Transformation und Selektion der Zellen als besonders geeignet herausgestellt.

Als Promotor für hph wird also bevorzugt p-gpdA, der Promotor der Glyceraldehyd-3-phosphatdehydrogenase aus *Aspergillus nidulans* genutzt.
20 Als Terminator für hph wird bevorzugt t-trpC, der Terminator des Gens trpC, codierend für Anthranilatsynthasekomponenten aus *Aspergillus nidulans* genutzt.

Als Vektoren haben sich Abkömmlinge des pBinAHyg Vectors als besonders geeignet herausgestellt. Der zur Transformation eingesetzte Vector umfasst also bevorzugterweise die SEQ ID NO: 3.

Hinzu kommen je nach gewünschtem Carotinoid oder dessen Vorstufe eine Sequenz codierend für eine Hydroxylase, Ketolase, Phytoendesaturase usw. wie diese zuvor beschrieben wurden. Die Vektoren umfassen
30 also in einer Ausführungsform der Erfindung die Sequenz SEQ ID NO: 69 codierend für die Phytoendesaturase. Die Vektoren umfassen ferner in

einer weiteren Ausführungsform der Erfindung die Sequenz SEQ ID NO: 72 codierend für eine Ketolase. Die Vektoren umfassen weiter in einer weiteren Ausführungsform der Erfindung die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydroxylase. Entsprechende Kombinationen
5 der zuvorgenannten Sequenzen liegen ebenso im Rahmen der Erfindung. So umfasst der Vector in einer Ausführungsform sowohl eine Sequenz SEQ ID NO: 72 codierend für eine Ketolase als auch die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydroxylase und ermöglicht so die Herstellung von Astaxanthin.

10

Insbesondere sind Vektoren ausgewählt aus der Gruppe bestehend aus den SEQ ID NO: 37 bis 51 und 62 im Rahmen der Erfindung einsetzbar.

15

Die genetisch veränderten Organismen können zur Produktion von Carotinoiden, Xanthophyllen oder deren Vorstufen, insbesondere Bixin, Phytoen, Astaxanthin, Zeaxanthin und Canthaxanthin verwendet werden. Auch können neue, im Wildtyp natürlicherweise nicht vorkommende Carotinoide durch Einbringung der entsprechenden genetischen Information von den gezielt genetisch veränderten Zellen bzw. dem durch sie gebildeten Mycel erzeugt und anschließend isoliert werden.

20

Die gentechnisch veränderten Zellen werden nach der Selektion kultiviert, so daß Carotinoiden oder deren Vorstufen bereitgestellt werden können.

25

Bevorzugterweise ist die Gewinnung von Carotinoiden oder deren Vorstufen mit den gezielt genetisch veränderten Zellen bzw. das durch sie gebildete Mycel möglich.

30

Die Kultivierung der Organismen unterliegt keinen Besonderheiten. Vorteilhafterweise werden, insbesondere bei der Verwendung von *Blakeslea*

trisporea, entgegengesetzte Paarungstypen gemeinsam kultiviert, da dies zu besserem Wachstum und Produktion führt.

5 Wird die gentechnische Veränderung nur in Zellen eines der vorkommen-
den Paarungstypen (bei *Blakeslea trispora* (+) oder (-)) durchgeführt, so
wird zur Kultivierung der entsprechend andere, nicht veränderte Paarungs-
typ zugesetzt, da so eine gute Produktion der Carotinoide oder deren Vor-
stufen aufgrund der von dem zweiten, nicht veränderten Paarungstyp ab-
gegebenen Substanzen (z. B. Trisporsäuren) zu erreichen ist. Vorteilhaft-
10 erweise wird jedoch die gentechnische Veränderung in Zellen beider Paa-
rungstypen vorgenommen und diese zusammen kultiviert. Hierdurch wird
ein besonders gutes Wachstum und eine optimale Produktion der Caroti-
noiden oder deren Vorstufen erreicht. Auch eine (künstliche) Zugabe der
Trisporsäuren ist möglich und sinnvoll.

15

Trisporsäuren sind Sexualhormone in Mucorales Pilzen, wie *Blakeslea*,
welche die Bildung von Zygomyceten und die Produktion von β -Carotin sti-
mulieren (van den Ende 1968, J. Bacteriol. 96:1298 - 1303, Austin et al.
1969, Nature 223:1178 - 1179, Reschke Tetrahedron Lett. 29:3435 -
20 3439, van den Ende 1970, J. Bacteriol. 101:423 - 428).

Es können alle dem Fachmann geläufigen Medien eingesetzt werden, so
weit sich diese zur Kultivierung der eingesetzten Organismen und deren
Carotinoid Produktion eignen. Insbesondere müssen bei Einsatz der
25 GVO keine Carotinoidbiosynthese Inhibitoren eingesetzt werden. Die ein-
gesetzten Medien beinhalten vorzugsweise Zusätze, wie eine oder mehre-
re Kohlenstoffquellen, eine oder mehrere Stickstoffquellen, Mineralsalze
und Thiamine. Bevorzugterweise werden Zusätze eingesetzt, wie sie aus
der WO 03/038064 A2, Seite 4, Zeile 30 bis Seite 5 Zeile 7 hervorgehen.
30 Besonders bevorzugt wird als Kohlenstoffquelle Glukose und als Stick-

stoffquelle Asparagin, pflanzliche oder tierische Extrakte, wie Baumwollsaatöl, Sojaöl, Baumwollsamemehl oder Hefe-Extrakt zugesetzt.

Die Kultivierung kann entweder unter aeroben oder anaeroben Bedingungen durchgeführt werden. Auch eine gemischte, zunächst aerobe und anschließend anaerobe Kultivierung, wie sie aus der DE 101 30 323 bekannt ist, ist möglich. Temperatur und Luftfeuchtigkeit werden dabei jeweils zum optimalen Wachstum eingestellt. Bevorzugterweise liegt die Temperatur bei der Kultivierung zwischen ca. 20 und ca. 34 °C, insbesondere zwischen ca. 26°C und ca. 28°C. Die Kultivierung kann ferner kontinuierlich, batch- oder satzweise erfolgen.

Die Kultivierung erfolgt vorzugsweise bis zu einem Feststoffgehalt zwischen etwa 1 und etwa 20 %, bevorzugt 3 und 15 % und besonders bevorzugt 4 und 11 %. Insbesondere ist wichtig dass die Kulturbrühe pumpbar bleibt, so dass sie in den nachfolgenden Verfahrensschritten ver- und bearbeitbar bleibt. Ist der Feststoffgehalt zu klein, so muss ein großer Aufwand bei der Aufkonzentrierung oder Trocknung betrieben werden.

Die Kultivierung bzw. Fermentation kann in den üblichen Apparaturen durchgeführt werden. Hierzu kommen alle für die jeweils eingesetzten Mikroorganismen und deren Produkte geeigneten Apparaturen in Betracht. Insbesondere solche, wie sie aus dem Römpp Lexikon Biotechnologie (1992 Georg Thieme Verlag, Stuttgart) unter dem Stichwort "Bioreaktor" auf Seiten 123 - 126 angegeben sind. Besonders bevorzugt ist der Einsatz von Rührkesselreaktoren mit versch. Einbauten, Blasensäulen verschiedener Bauarten, etc.

Die nach dem erfindungsgemäßen Verfahren bereitgestellten Carotinoiden oder deren Vorstufen, insbesondere Bixin, Phytoen oder Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin eignen sich besonders

zur Herstellung von Zusätzen für Futter-, Nahrungs- und Nahrungsergänzungsmittel, kosmetischen, pharmazeutischen oder dermatologischen Zubereitungen.

- 5 Die Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe aus der Kultur der gentechnisch veränderten Mikroorganismen erfolgt nach zwei Varianten a) oder b), bevorzugt ist auch eine Kombination aus a) und b);

10

a:

- I) Abtrennung der Biomasse,
 - 15 IA) ggf. Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser,
 - IB) Sterilisation und Zellaufschluß der Biomasse,
 - IC) ggf. Trocknung und/oder homogene Verteilung und
- II) partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse,
 - 20 IIA)
 - 1) Entfernung von Lösemittelresten aus der Carotinoidhaltigen Biomasse,
 - 2) ggf. homogene Suspension der Biomasse mit einem Biomasse-Feststoffgehalt > 2 % und < 50 %, und
 - 25 3) Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels,
 - IIB)
 - 1) Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration;

30

oder b):

- 5 I) Homogene Suspendierung der Feststoffe der Kulturbrühe
und
- IIA) bei einem Feststoffgehalt der Kulturbrühe von > 2 %
 1) ggf. Konzentration der Kulturbrühe auf einen Fest-
 stoffgehalt < 50 % und
10 2) Trocknung der Kulturbrühe zur Herstellung des Nah-
 rungsmittels
oder
- IIB) bei einem Feststoffgehalt von < 2 % der Kulturbrühe,
15 1) Konzentration der Kulturbrühe auf einen Feststoffge-
 halt > 2 % und < 50 % und
 2) Trocknung der Suspension zur Herstellung des Nah-
 rungsmittels,
oder
- 20 IIC) unabhängig vom Feststoffgehalt der Kulturbrühe,
 1) Abtrennung der Biomasse,
 2) ggf. Waschen der Biomasse mit Carotinoide nicht lö-
 senden Lösungsmitteln, insbesondere Wasser,
25 3) Sterilisation und Zellaufschluß,
 4) ggf. Trocknung und homogene Verteilung,
 5) partielle Extraktion der Carotinoide aus der Biomasse
 mittels eines Carotinoide lösendes Lösungsmittels,
 5a) Abtrennung der Carotinoid-haltigen Biomasse vom
30 Carotinoid-haltigen Lösungsmittel,

- 5b) Entfernung von Lösemittelresten aus der Biomasse und
- 5c) Trocknung der Biomasse zur Herstellung des Nahrungsmittels,
- 5 6) Kristallisation der Carotinoide aus dem in 5a) verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.

10 Die erfindungsgemäße Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe aus der Kultur der gentechnisch veränderten Mikroorganismen erfolgt nach zwei Varianten a) oder b) ermöglicht die gleichzeitige Herstellung von zwei Produkten.

15

Durch die erfindungsgemäße Kombination der Herstellung von zwei Produkten, insbesondere bei der Bereitstellung gemäß Variante a), nämlich dem mindestens einem Carotinoid und dem Carotinoid-haltigem Nahrungsmittel ist keine vollständige Extraktion der Carotinoide aus der Bio-

20 masse nötig, so dass der Aufwand bei der Extraktion geringer ausfällt. Das Carotinoid muss trotz vollständiger Verwertung nur partiell extrahiert werden, ohne dass es zu Produktverlusten kommt. Dies bedingt geringere Lösungsmittelmengen und damit einhergehend einen geringeren Aufwand bei den Maßnahmen zu deren Wiederverwendung. Zudem werden Abfälle

25 weitestgehend vermieden, da die Biomasse nicht als Abfall anfällt, sondern zum hochwertigen Nahrungsmittel weiterverarbeitet wird. Somit ergeben sich geringere Kosten für die Verfahren durch Ausnutzen von Synergien.

30 Nach dem erfindungsgemäßen Verfahren mit Bereitstellung gemäß Variante b) erhältliche Nahrungsmittel enthalten also bereits nach der Herstel-

lung große Mengen an Carotinoiden, die nicht zugesetzt werden müssen. Dadurch, dass das Nahrungsmittel neben dem mindestens einen Carotinoid auch *Blakeslea trispora* enthält ist sein Nährstoffgehalt zudem gesteigert. Insbesondere ist der Nährstoffgehalt nach den bevorzugten Alternativen IIA und IIB stark gesteigert, da es neben dem mindestens einen Carotinoid und *Blakeslea trispora* zusätzlich alle Medienbestandteile der Fermentation enthält. Ferner benötigt das Verfahren keine zusätzlichen, aufwendigen Aufarbeitungs- und Herstellungsschritte, sondern die homogenisierte und ggf. entwässerte *Blakeslea trispora*-haltige Kulturbrühe kann direkt ohne Umwege zur Herstellung des Nahrungsmittels getrocknet werden. Es fallen demnach praktisch keine Abfälle an, abgesehen vom wässrigem Medium bei der Alternative IIB, welches jedoch unproblematisch in einer Kläranlage gereinigt werden kann. Zusätzlich wird in allen drei Alternativen die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da gemäß IIA und IIB keine verlustreichen Trenn- bzw. Aufarbeitungsschritte vorgenommen werden müssen. Bei der Alternative IIIC wird ebenfalls die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da ein Teil in der Biomasse zum Nahrungsmittel verarbeitet wird und der andere Teil zur Gewinnung reiner Carotinoide extrahiert wird. Durch die erfindungsgemäße Kombination der Herstellung von zwei Produkten gemäß IIC, nämlich dem Carotinoid-haltigem Nahrungsmittel und den Carotinoiden an sich, ist vorteilhaft, dass wiederum im wesentlichen keine Abfälle entstehen und eine vollständige Extraktion der Carotinoide aus der Biomasse unnötig ist, so dass sonst bei der Extraktion anfallende Aufwand geringer ausfällt. Das oder die wertvollen Carotinoide müssen trotz vollständiger Verwertung nur partiell extrahiert werden, ohne dass es zu Produktverlusten kommt. Dies bedingt geringere Lösungsmittelmengen und damit einhergehend einen geringeren Aufwand bei den Maßnahmen zu deren Wiederverwendung. Zudem werden Abfälle weitestgehend vermieden, da die Biomasse nicht als Abfall anfällt, sondern zum hochwertigen Nahrungsmittel verarbeitet

wird. Somit ergeben sich geringere Kosten für die Verfahren durch Ausnutzen von Synergien.

Unter "hochrein" soll in der vorliegenden Anmeldung eine Reinheit des
5 mindestens einen Carotinoids von mindestens 95%, bevorzugt > 95%, vorzugsweise > 96%, besonders bevorzugt > 97%, ganz besonders bevorzugt > 98%, höchst bevorzugt > 99% verstanden werden.

Als nach dem erfindungsgemäßen Verfahren herstellbare Carotinoide
10 kommen alle natürlichen und künstlichen Carotine und Xanthophylle in Betracht. Insbesondere ist das mindestens eine Carotinoid aus der Gruppe bestehend aus Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, Lutein, Phytofluen, Bixin und
15 Phytoen ausgewählt. Bevorzugterweise handelt es sich um Astaxanthin oder Zeaxanthin. Die Carotinoide können nach dem erfindungsgemäßen Verfahren einzeln oder als Gemische von zwei oder mehrerer der zuvor genannten Carotinoide erhalten werden. Insbesondere bei Einsatz der weiter unten angegebenen gentechnisch veränderten Organismen (GVO)
20 kann das oder können die Carotinoide gezielt hergestellt werden.

Als Nahrungsmittel werden Zusammensetzungen angesehen, die der Ernährung dienen. Darunter fallen auch Zusammensetzungen für die Ergänzung der Ernährung. Insbesondere werden als Nahrungsmittel Tierfuttermittel und Tierfutterergänzungsmittel angesehen.
25

Nach der Kultivierung kann, gemäß Variante a) der Bereitstellung, die Biomasse von der Kulturbrühe abgetrennt. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden zur
30 fest/flüssig-Trennung eingesetzt werden. Hierunter fallen insbesondere die mechanischen Verfahren, wie Filtration und Zentrifugation, die auf der

Ausnutzung von Schwerkraft, Zentrifugalkraft, Druck oder Vakuum beruhen. Zu den einsetzbaren Verfahren und Apparaten gehören daneben u. a. Querstromfiltration bzw. Mebrantechniken wie Osmose, umgekehrte Osmose, Mikrofiltration, Ultrafiltration, Nanofiltration, Kuchenfiltrationsverfahren (z.B. mittels Pressfilterautomaten, (Membran-, Rahmen- oder Kammer-)Filterpressen, (Rühr-)drucknutschen, Saugnutschen, (Vakuum-)bandfiltern, (Vakuum-)Trommelfiltern, Drehfiltern, Kerzenfiltern), Zentrifugationsverfahren mittels kontinuierlich oder diskontinuierlich betriebener Zentrifugen oder Filterzentrifugen (z.B. Stülpfilterzentrifugen, Schälzentrifugen, Schubzentrifugen, Siebschneckenzentrifugen, Gleitzentrifugen, Separatoren oder Dekantern), Verfahren unter Ausnutzung der Schwerkraft wie Flotation, Sedimentation, Sink-Schwimm-Aufbereitung und Klären. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekanterns oder durch Filtration, mittels einer Mebranfiltrationseinheit durchgeführt.

In dem zweiten Schritt der Bereitstellung nach Variante b) wird eine homogen verteilte Suspension der Feststoffe in der Kulturbrühe erzeugt. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden verwendet werden. Insbesondere kommen dazu Dispergiergeräte, wie ein Ultra-Turrax® (im Labormaßstab) zum Einsatz. Ein Zellaufschluss ist nicht notwendig, kann aber vorgenommen werden.

Falls notwendig, kann die Kulturbrühe entwässert werden, um einen geeigneten Feststoffgehalt zwischen $> 2 \%$ und $< 50 \%$ zu erreichen. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden zur fest/flüssig-Trennung eingesetzt werden. Hierunter fallen insbesondere die mechanischen Verfahren, wie Filtration und Zentrifugation, die auf der Ausnutzung von Schwerkraft, Zentrifugalkraft, Druck oder Vakuum beruhen. Zu den einsetzbaren Verfahren und Apparaten gehören daneben u. a. Querstromfiltration bzw. Mebrantechniken wie Osmose,

umgekehrte Osmose, Mikrofiltration, Ultrafiltration, Nanofiltration, Kuchenfiltrationsverfahren (z.B. mittels Pressfilterautomaten, (Membran-, Rahmen- oder Kammer-)Filterpressen, (Rühr-)drucknutschen, Saugnutschen, (Vakuum-)bandfiltern, (Vakuum-)Trommelfiltern, Drehfiltern, Kerzenfiltern),
5 Zentrifugationsverfahren mittels kontinuierlich oder diskontinuierlich betriebener Zentrifugen oder Filterzentrifugen (z.B. Stülpfilterzentrifugen, Schälzentrifugen, Schubzentrifugen, Siebschneckenzentrifugen, Gleitzentrifugen, Separatoren oder Dekantern), Verfahren unter Ausnutzung der Schwerkraft wie Flotation, Sedimentation, Sink-Schwimm-Aufbereitung
10 und Klären. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekantern oder wird durch Filtration, mittels einer Membranfiltrationseinheit durchgeführt. Anschließend wird die Kulturbrühe getrocknet. Hierzu können wiederum alle dem Fachmann bekannten Verfahren und Apparate eingesetzt werden.
15 Insbesondere eignen sich Apparate zur thermischen Trocknung wie Konvektions-, Kontakt- und Strahlungstrocknung, z.B. Horden-, Kammer-, Kanal-, Flachbahn-, Teller-, Drehtrommel-, Rieselschacht-, Siebband-, Strom-, , Wirbelschicht-, Fließbett-, Schaufel-, Kugelbett-, , Heizteller-, Dünnschicht-, Walzen-, Band-, Siebtrommel-, Schnecken-, Taumel-, Kontakt-Scheiben-, Infrarot-, Mikrowellen- und Gefriertrockner, Sprühtrockner
20 oder Sprühtrockner mit integrierter Wirbelschicht, die durch Dampf, Öl, Gas oder elektrischen Strom ggf. beheizt und ggf. unter Vakuum betrieben werden. Die Betriebsweise kann dabei je nach Apparat kontinuierlich oder diskontinuierlich sein. Daneben oder damit in Kombination können die
25 oben bereits angegeben mechanischen Verfahren zur Fest/flüssig-Trennung verwendet werden.

Eine Granulierung durch Extrusion wie dies aus der WO 97/36996 A2 hervorgeht ist jedoch nicht notwendig. Durch die Trocknung wird das Nahrungsmittel haltbar und lagerfähig.

Insbesondere wird die Kulturbrühe sprühgetrocknet. Bevorzugt wird zur Trocknung die Sprühtrocknung eingesetzt, wie sie aus der DE 101 04 494 A1, DE-A-12 11 911 oder EP 0 410 236 A1 bekannt sind. Ergänzend wird auf vgl. Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Sprühtrocknung" und Römpp Lexikon Biotechnologie, Georg Thieme Verlag, 1992, "Zerstäubungstrocknung" verwiesen. Die Sprühtrocknung bietet den Vorteil der kurzen Verweilzeit des Produkts in der heißen Zone des Trockners, so dass eine besonders schonende Trocknung erzielt wird.

10

Bei der Sprühtrocknung werden Eingangstemperaturen von ca. 115°C – 180°C, bevorzugt 120°C – 130°C, und Ausgangstemperaturen von ca. 50°C – 80°C, bevorzugt 55°C – 70°C gewählt. Als Trocknungsgas wird vorzugsweise Stickstoff eingesetzt.

15

Gegebenenfalls können zur Erzielung einer besseren Rieselfähigkeit Rieselfähigungsmittel, wie Kieselsäuren etc. zugesetzt werden. Der Einsatz von inerten Trägermaterialien, d.h. niedermolekularen anorganischen Trägern wie NaCl, CaCO₃, Na₂SO₄ oder MgSO₄, organischen Trägern wie Glucose, Fructose, Saccharose, Dextrine oder Stärkeprodukten (Roggen-, Gersten-, Hafermehl, Weizengrießkleie) ist denkbar.

20

Das getrocknete Produkt weist bevorzugt eine Restfeuchte von weniger als 10 %, bevorzugt weniger als 5 % bezogen auf die Trockenmasse auf. Sein Carotinoidgehalt liegt zwischen 0,05 und 20 %, insbesondere 1 und 10 % bezogen auf die Trockenmasse.

25

Das so hergestellte Nahrungsmittel kann entweder direkt verwendet werden oder mittels weiterer Zusätze aufbereitet werden, so wie dies ebenfalls aus der DE 101 04 494 A1 bekannt ist.

30

Gemäß der Alternative IIC wird nach der Kultivierung und vor der Trocknung der Biomasse, die Biomasse von der Kulturbrühe zunächst abgetrennt. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden zur fest/flüssig-Trennung eingesetzt werden, wie
5 sie bereits oben bei der Entwässerung genannt wurden. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekanters oder durch Membranfiltration durchgeführt.

Anschließend erfolgt optional das Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser, wodurch insbesondere Wasser lösliche Komponenten entfernt werden. Dieser Schritt
10 kann gegebenenfalls unter Verwendung weiterer Carotinoide nicht lösenden Lösungsmittel (z. B. Alkohole) ergänzt werden, was aber im Rahmen der Erfindung nicht notwendig ist und zur Vermeidung von Abfällen nicht
15 bevorzugt ist.

Anschließend erfolgen die Sterilisation und der sich anschließende oder gleichzeitige Zellaufschluß der Zellen in der Biomasse. Durch die Sterilisation werden die Mikroorganismen abgetötet und gegebenenfalls vorhandene Enzymaktivität beendet. Dies ist zur Verhinderung des Abbaus der
20 Biomasse bzw. der darin enthaltenen Stoffe, insbesondere der Carotinoide und für die Haltbarkeit von Bedeutung.

Die Sterilisation kann mit einem üblichen, dem Fachmann geläufigen Verfahren durchgeführt werden. Hierzu gehören die Sterilisation mittels
25 Dampf, insbesondere bei Temperaturen größer 120 °C unter Druck (≥ 1 bar) und Zeitdauern von \geq ca. 20 min. sowie die Behandlung mit energiereichen Strahlen, wie UV-, Mikrowellen, Gamma- oder Beta-Strahlen. Bevorzugterweise erfolgt die Sterilisation im Rahmen des erfindungsgemäßen Verfahrens mittels Dampf oder Mikrowellenstrahlung.
30

Durch den nachfolgenden oder gleichzeitigen Zellaufschluß, werden die innerhalb der Zellen vorliegenden Carotinoide freigesetzt. Der Zellaufschluß kann ebenfalls mit allen dem Fachmann bekannten üblichen Verfahren erfolgen. Hierzu gehören mechanische und nicht mechanische Methoden. Zu den mechanischen Methoden zählen Trockenmahlen, Naßmahlen, Rühren, Homogenisieren (z.B. im Hochdruckhomogenisator) und die Verwendung von Ultraschall oder Mikrowellen. Als nicht mechanische Methoden kommen physikalische, chemische und biochemische Methoden in Betracht. Hierzu gehören Kurzzeiterhitzen, Kurzzeitgefrieren, Osmotischer Schock, Trocknung, Behandlung mit Säuren oder Laugen sowie ein enzymatischer Aufschluss. Günstiger Weise wird zum Zellaufschluss jedoch das zur Sterilisation verwendete Verfahren eingesetzt. Bevorzugterweise wird also ebenfalls der Zellaufschluss mittels Dampf oder Mikrowellen Strahlung durchgeführt.

15

Die Sterilisation und/oder der Zellaufschluss können kontinuierlich oder diskontinuierlich durchgeführt werden.

20

Die Sterilisation und/oder der Zellaufschluss können im zur Kultivierung eingesetzten Bioreaktor oder in anderen Apparaturen, wie Autoklaven usw. durchgeführt werden. Bei kontinuierlicher Durchführung kann das aus der WO 01/83437 A1 bekannte Mikrowellen verwendende Verfahren und entsprechende Apparaturen eingesetzt werden.

25

Vor der Extraktion wird die Biomasse gegebenenfalls getrocknet und/oder homogenisiert. Hierzu können wiederum alle dem Fachmann bekannten üblichen Verfahren und Geräte eingesetzt werden. Insbesondere eignen sich Apparate zur thermischen Trocknung wie Konvektions-, Kontakt- und Strahlungstrocknung, z.B. Horden-, Kammer-, Kanal-, Flachbahn-, Teller-, Drehtrommel-, Rieselschacht-, Siebband-, Strom-, , Wirbelschicht-, Fließbett-, Schaufel-, Kugelbett-, , Heizteller-, Dünnschicht-, Walzen-, Band-,

30

Siebtrommel-, Schnecken-, Taumel-, Kontakt-Scheiben-, Infrarot-, Mikrowellen- und Gefriertrockner, Sprühtrockner oder Sprühtrockner mit integrierter Wirbelschicht, die durch Dampf, Öl, Gas oder elektrischen Strom ggf. beheizt und ggf. unter Vakuum betrieben werden. Die Betriebsweise
5 kann dabei je nach Apparat kontinuierlich oder diskontinuierlich sein. Daneben oder damit in Kombination können die oben bereits angegeben mechanischen Verfahren zur Fest/flüssig-Trennung verwendet werden.

Eine Granulierung durch Extrusion wie dies aus der WO 97/36996 A2 hervorgeht ist jedoch nicht notwendig.
10

Anschließend erfolgt die partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse. Sowohl in dem
15 Lösungsmittel als auch in der Biomasse sind nun Carotinoide enthalten, wobei sich in dem Lösungsmittel bevorzugterweise der Großteil der Carotinoide befindet.

Aus dem Lösungsmittel werden anschließend die hochreinen Carotinoide
20 isoliert, wohingegen die Biomasse zu einem hochwertigen, Carotinoidhaltigen Nahrungsmittel weiterverarbeitet wird, welches durch den vorhergehenden Zellaufschluss auch eine gute Bioverfügbarkeit der Carotinoide aufweist.

25 Unter partieller Extraktion soll demnach die bewusst unvollständige Extraktion der Carotinoiden aus der Biomasse verstanden werden (vgl. oben). Bevorzugterweise wird im Rahmen der Erfindung durch die Extraktion also weniger als 100 % der in der Biomasse enthaltenen Gesamtmenge der Carotinoide aus dieser extrahiert.

30 Dies ist von großem Vorteil, da der Aufwand zur Extraktion mit der abnehmenden Menge Carotinoid in der Biomasse überproportional zunimmt.

Zur Extraktion werden Lösungsmittel eingesetzt, die Carotinoide lösen, wie z. B. Hexan, Ethylacetat, Dichlormethan oder überkritisches Kohlendioxid. Bevorzugterweise wird erfindungsgemäß als Lösungsmittel Dichlormethan
5 oder überkritisches Kohlendioxid eingesetzt, wobei beim Einsatz von überkritischem Kohlendioxid die darin enthaltenen Carotinoide anschließend in Dichlormethan überführt werden können oder das Wertprodukt direkt durch Entspannung des Kohlendioxids gewonnen werden kann. Dabei werden die Mengen der Lösungsmittel und Durchmischungszeiten
10 derart gewählt, dass die gewünschte Menge Carotinoide aus der Biomasse extrahiert wird. Insbesondere wird der Extraktionsschritt nur einmal durchgeführt, was technisch und wirtschaftlich sinnvoll ist (vgl. oben).

Zur Durchführung der Extraktion können alle üblichen Verfahren und Apparaturen eingesetzt werden. Insbesondere wird bei nicht getrockneter,
15 aber aufgeschlossener Biomasse eine flüssig/flüssig (Carotinoid liegt in flüssigen Zellbestandteilen gelöst vor und wird daraus extrahiert und bei getrockneter Biomasse eine fest/flüssig Extraktion durchgeführt. Es können Kalt- und Heißeextraktion in bestimmten Temperaturbereichen, sowohl
20 kontinuierliche (z.B. Soxhlet-Extraktion, Perforation und Perkolation) als auch diskontinuierliche Verfahren, zu denen beispielsweise Ausschütteln, Auslaugen, Auskochen und Digerieren gehören, verwendet werden. Sie können auch im Gegenstromverfahren durchgeführt werden.

25 Für die flüssig/flüssig Extraktion können beispielsweise Blasensäulen, , pulsierende Kolonnen, Kolonnen mit rotierenden Einbauten, Mixer-Settler-Batterien oder Rührkessel usw. verwendet werden.

Die fest/flüssig Extraktion kann mittels üblicher Apparaturen durchgeführt
30 werden. Vorzugsweise werden Rührkessel oder Mixer-Settler-Apparate eingesetzt.

Alternativ kann der Zellaufschluß ohne vorherige Abtrennung des Fermentationsmediums erfolgen und sich dann eine direkte Trennung einer sich bildenden Carotinoidsuspension von der Biomasse z. B. mittels eines Dekan-
5 ters durchgeführt werden. Anschließend wird die Carotinoidsuspension in Dichlormethan aufgenommen und weiterverarbeitet oder alternativ durch Wäschen mit verschiedenen wässrigen Lösungen aufgereinigt.

10 Zur Isolierung der hochreinen Carotinoide aus dem Lösungsmittel wird eine Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration durchgeführt. Die verbleibende Mutterlauge kann nach Destillation dem Verfahren erneut zugeführt werden, so dass Produktverluste trotz geringem Aufwand minimiert werden.

15 Die Kristallisation kann wie üblich erfolgen. Ergänzend wird auf vgl. Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Kristallisation" verwiesen.

20 Bevorzugterweise erfolgt die Kristallisation durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel. Es wird also kontinuierlich die Löslichkeit der Carotinoide erniedrigt, bis diese als reine Kristalle ausfallen. Hierbei wird vorzugsweise ein "niederer Alkohol" oder Wasser verwendet. Als niederer Alkohol werden aliphatische Alkohole mit 1 bis 4 Kohlenstoffatomen angesehen. Hierzu gehören Methanol,
25 Ethanol, Propanol, Isopropanol, 1-Butanol, tert.-Butanol und sec.-Butanol. Bevorzugterweise wird Methanol eingesetzt.

30 Die Carotinoid-Lösung kann dabei erwärmt werden, wobei die Temperatur vorzugsweise < 100 °C, insbesondere < 60 °C gehalten wird, so dass Dichlormethan abdestilliert wird. Auch der Einsatz von Vakuum ist denk-

bar. Anschließend werden die Carotinoidkristalle isoliert, welches durch übliche Maßnahmen, insbesondere durch Filtration erfolgen kann. Es können sich, falls gewünscht, weitere optionale Trocknungs- und/oder Reinigungsschritte anschließen. Notwendig sind diese jedoch nicht, da die Carotinoid Kristalle bereits hochrein sind.

Die Carotinoide fallen als hochreine Kristalle an und weisen eine Reinheit von mindestens 95%, bevorzugt > 95%, vorzugsweise > 96%, besonders bevorzugt > 97%, ganz besonders bevorzugt > 98%, höchst bevorzugt > 99% auf.

Die erzielbaren Ausbeuten liegen zwischen 45% und 95%, bevorzugt zwischen 70% und 95% bezogen auf die in der Kulturbrühe vorliegende Menge (0,5 – 15 g/L, bevorzugt 1 – 10 g/L).

Zur Weiterverarbeitung der ebenfalls Carotinoid-haltigen Biomasse zu einem hochwertigen Nahrungsmittel wird zunächst eine Entfernung von Lösemittelresten aus der Carotinoid-haltigen Biomasse vorgenommen. Hierzu erfolgt bevorzugterweise eine Wasserdampfdestillationen bzw. ein so genanntes Strippen mit Wasserdampf (vgl. Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Strippen").

Danach kann gegebenenfalls die Biomasse in der oben abgetrennten Kulturbrühe homogen suspendiert werden, wobei ein Feststoffgehalt > 100 g/L und < 600 g/L eingehalten werden sollte, so daß die nachfolgende Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels ohne technische Schwierigkeiten erfolgen kann. D.h. die Suspension muß pumpbar sein. Als Trocknungsverfahren kommen alle bereits genannten Verfahren und Apparaturen in Frage. Insbesondere wird zur Trocknung die Sprühtrocknung eingesetzt. Dabei kann wie aus der DE 101 04 494 A1 bekannt verfahren werden.

Bei der Sprühtrocknung werden Eingangstemperaturen von ca. 100°C – 180°C, bevorzugt 120°C – 130°C, und Ausgangstemperaturen von ca. 50 – 80°C, bevorzugt 55°C – 70°C gewählt. Als Trocknungsgas wird vorzugsweise Stickstoff eingesetzt.

Das so hergestellte Nahrungsmittel kann entweder direkt verwendet werden oder mittels weiterer Zusätze aufbereitet werden, so wie dies ebenfalls aus der DE 101 04 494 A1 bekannt ist.

10

Als Nahrungsmittel werden Zusammensetzungen angesehen, die der Ernährung dienen. Darunter fallen auch Zusammensetzungen für die Ergänzung der Ernährung. Insbesondere werden als Nahrungsmittel Tierfuttermittel und Tierfutterergänzungsmittel angesehen. Ergänzend wird auf Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Nahrungsmittel" verwiesen.

15

Das trockene Produkt weist bevorzugt eine Restfeuchte von weniger als 5 % bezogen auf die Trockenmasse auf. Sein Carotinoidgehalt liegt zwischen 0,05 und 20 %, insbesondere 1 und 10 % bezogen auf die Trockenmasse. Der gewünschte Carotinoidgehalt ist über das Ausmaß der Extraktion steuerbar (vgl. oben).

20

Nach dem erfindungsgemäßen Verfahren erhältliche Nahrungsmittel enthalten also bereits nach der Herstellung große Mengen an Carotinoiden, die nicht zugesetzt werden müssen. Dadurch, dass das Nahrungsmittel neben dem mindestens einen Carotinoid auch Biomasse enthält ist sein Nährstoffgehalt zudem gesteigert. Insbesondere ist der Nährstoffgehalt nach der bevorzugten Alternative stark gesteigert, in dem es neben dem mindestens einen Carotinoid und Biomasse zusätzlich alle Medienbestandteile der Fermentation enthält. Es fallen demnach praktisch keine

25

30

Abfälle an, abgesehen von wässrigen Medien, welche jedoch unproblematisch in einer Kläranlage gereinigt werden können. Zusätzlich wird die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da keine verlustreichen Trenn- bzw. Aufarbeitungsschritte vorgenommen werden müssen, um die gesamte Menge Carotinoid zu extrahieren.

Die in dem oben beschriebenen, erfindungsgemäßen Verfahren eingesetzten Lösungsmittel werden alle soweit wie möglich aufbereitet und anschließend wieder verwendet bzw. dem Verfahren erneut zugeführt. Insbesondere wird das eingesetzte Dichlormethan bereits beim Lösungsmittel-
austausch gereinigt und steht anschließend zur erneuten Verwendung bereit. Der niedere Alkohol bzw. das Methanol wird z. B. destillativ gereinigt und ebenfalls wieder verwendet. Als Abfall fallen lediglich der Destillationssumpf an, der zusammen mit den wässrigen Medien gefahrlos einer Kläranlage zugeführt werden kann, wo letztendlich nur eine geringe Menge Klärschlamm als tatsächlicher Abfall anfällt. Somit ist das beschriebene Verfahren im wesentlichen abfallfrei.

20

Die Erfindung wird nachfolgend an Hand von Beispielen näher ausgeführt.

25 **A) Kultivierung von Blakeslea trispora**

Folgende Medien wurden zur Fermentation von Blakeslea trispora zur Produktion der Carotinoide eingesetzt:

Medium 1:

	Glucose	10,00 g/l
30	Baumwollsaatoel	30,00 g/l
	Sojaoel	30,00 g/l

	Dextrin	60,00 g/l
	Baumwollsamemehl	75,00 g/l
	Triton X 100	1,20 g/l
	Ascorbinsäure	6,00 g/l
5	Milchsäure	2,00 g/l
	KH_2PO_4	0,50 g/l
	$\text{MnSO}_4 \times \text{H}_2\text{O}$	100 mg/l
	Thiamin-HCl	2 mg/l
	Isoniazid (Isonicotinsäurehydrazid)	0,75 g/l
10	Der pH wurde auf 6,5 eingestellt.	

Medium 2:

	Glucose	20 g/l
	Asparagin	2,00 g/l
15	KH_2PO_4	5,00 g/l
	$\text{MgSO}_4 \times 7 \text{ H}_2\text{O}$	0,50 g/l
	CaCl_2	28 mg/l
	Thiamin-HCl	1,00 mg/l
	Citronensäure	2,00 mg/l
20	$\text{Fe}(\text{NO}_3)_3 \times 9 \text{ H}_2\text{O}$	1,50 mg/l
	$\text{ZnSO}_4 \times 7 \text{ H}_2\text{O}$	1,00 mg/l
	$\text{MnSO}_4 \times \text{H}_2\text{O}$	0,30 mg/l
	$\text{CuSO}_4 \times 5 \text{ H}_2\text{O}$	0,05 mg/l
	$\text{Na}_2\text{MoO}_4 \times 2 \text{ H}_2\text{O}$	0,05 mg/l
25		

Medium 3

	Glucose	70,00 g/l
30	Asparagin	2,00 g/l
	Hefe Extrakt	1,00 g/l

KH ₂ PO ₄	1,50 g/l
MgSO ₄ x 7 H ₂ O	0,50 g/l
Span 20	1,00 g/l
Thiamin-HCl	5,0 mg/l

- 5 Der pH wurde auf 5,5 eingestellt.

Mit Sporensuspensionen von *Blakeslea trispora* ATCC 14272 Mating Type (-) die 10⁸ (für Medium 2) bzw. 10⁷ (für Medium 1 und 3) Sporen enthielten, wurden je 200 ml der beschriebenen Medien angeimpft. Die Kultivierung erfolgte jeweils in 1-l-Erlenmeyerkolben mit Schikanen. Mit jedem Medium wurden sechs identische Kolben angesetzt und über 7 Tage bei 28°C und 140 UpM im Schüttler inkubiert.

B) Gentechnische Veränderung von *Blakeslea Trispora*

15

Material und Methoden

Molekulargenetische Arbeiten wurden, wenn nicht anders beschrieben, nach den Methoden in Current Protocols in Molecular Biology (Ausubel et al., 1999, John Wiley & Sons) durchgeführt.

20

Stämme und Wachstumsbedingungen

Die *Blakeslea trispora* Stämme ATCC 14271 (Paarungstyp(+)) und ATCC14272 (-) (ein Wildtyp) wurde erhalten. Paarungstyp (-) wurden von der American Type Culture Collection. erhalten. Die Anzucht von *B. trispora* erfolgte in MEP-Medium (Malzextrakt-Pepton-Medium): 30 g/l Malzextrakt (Difco), 3 g/l Pepton (Soytone, Difco), 20 g/l Agar, Einstellung pH 5,5, ad 1000 ml mit H₂O bei 28 °C.

Die Anzucht von *Agrobacterium tumefaciens* LBA4404 erfolgte nach Hoekema et al. (1983, Nature 303:179-180) bei 28 °C für 24. h in Agrobacterien-Minimal Medium (AMM): 10 mM K₂HPO₄, 10 mM KH₂PO₄, 10 mM Glu-

cose, MM-Salze (2,5 mM NaCl, 2 mM MgSO₄, 700 µM CaCl₂, 9 µM Fe-SO₄, 4 mM (NH₄)₂SO₄).

Transformation von *Agrobacterium tumefaciens*

- 5 Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Transformation von *Blakeslea trispora*

- 15 Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD₆₀₀ von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₀₀ von ca. 0,6 angezogen.
- 20 Zur Co-Inkubation von *Blakeslea* ATCC 14271 bzw. ATCC14272 und *Agrobacterium* wurden 100 µl Agrobakteriensuspension mit 100 µl *Blakeslea* Sporensuspension (10⁷ Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die
- 25 Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt. Zur Selektion auf transformierte *Blakesleazellen* enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer
- 30 von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-

Asparagin, 50 mg/l $\text{MgSO}_4 \times 7\text{H}_2\text{O}$, 150 mg/l KH_2PO_4 , 25 $\mu\text{g/l}$ ThiaminHCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, 100 mg/L Hygromycin, 100 mg/L Cefotaxim, pH 5,5, 18 g/l Agar) ausplattiert. Zur Isolierung einzelner gentechnisch veränderter Sporen wurden die Sporen durch ein
5 FACS Gerät der Fa. BectonDickson (Modell Vantage+Diva Option) einzeln auf Selektivmedium abgelegt.

Mutagenese mit MNNG

Zur Reduzierung der Anzahl von Kernen pro Spore wurde eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt. Hierfür wurde zunächst eine Sporensuspension mit 1×10^7 Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 $\mu\text{g/ml}$ zugegeben. Die Zeit der Inkubation in MNNG wurde so gewählt, dass die
10 Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und plattiert.
15

Selektion homonukleater Zellen

20 Die Selektion homonukleater Zellen von *Blakeslea trispora* carB^- erfolgte analog zum Versuchsprotokoll für *Phycomyces blakesleeanus* (Roncero et al., 1984, Mutation Research, 125:195-204), modifiziert durch Wachstum in Gegenwart von 5-Carbon-5-Deazariboflavin (1 $\mu\text{g/ml}$) und Hygromycin 100 ($\mu\text{g/ml}$).

25

Herstellung genetisch veränderter *Blakeslea trispora* durch Agrobacterium-vermittelte Transformation

Herstellung des rekombinanten Plasmids pBinAHyg

Aus dem Plasmid pANsCos1 (Fig.1, Osiewacz, 1994, Curr. Genet. 26:87-90, SEQ ID NO: 4) wurde die *gpdA*-*hph*-*trpC*-Kassette als *Bgl*III/*Hind*III Fragment isoliert und in das mit *Bam*HI/*Hind*III geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet (Fig. 2, SEQ ID NO: 3) und enthielt das *E. coli* Hygromycin-Resistenzgen (*hph*) unter Kontrolle des *gpd* Promotors (SEQ ID NO: 1) und des *trpC* Terminators (SEQ ID NO: 2) aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind. Die in den weiter unten beschriebenen Ausführungsbeispielen genannten Vektoren sind Abkömmlinge von pBinAHyg.

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in *Agrobacterium tumefaciens*

Nachfolgend wird beispielhaft die Übertragung des Plasmids pBinAHyg in *Agrobacterien* beschrieben. Die Übertragung der Abkömmlinge erfolgte analog.

Das Plasmid pBinAHyg wurde in den *Agrobakterienstamm* LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der *Agrobakterienanzucht* folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora*

Zur Transformation wurden die *Agrobakterien* nach 24 h Anzucht in AMM auf eine OD₆₆₀ von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES

(pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₆₀ von ca. 0,6 angezogen.

- 5 Zur Co-Inkubation von *Blakeslea trispora* (B.t.) und *Agrobacterium tumefaciens* (A.t.) wurden 100 µl Agrobakteriensuspension mit 100 µl Blakeslea Sporensuspension (10^7 Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt.

- Zur Selektion auf transformierte Blakeslea-Zellen enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen
- 15 Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l MgSO₄ x 7H₂O, 150 mg/l KH₂PO₄, 25 µg/l Thiamin-HCl, 100 mg/l Yeast Extract,
- 20 100 mg/l Na-desoxycholat, pH 5,5, 100 mg/l Cefotaxim, 100 mg/l Hygromycin, 18 g/l Agar) ausplattiert. Die Übertragung von Sporen auf frische Selektionsplatten wurde dreimal wiederholt. Auf diese Weise wurde die Transformante *Blakeslea trispora* GVO 3005 isoliert. Alternativ erfolgte zur Selektion der GVO (gentechnisch veränderten Organismen) die Einzelab-
- 25 lage der Sporen durch den BectonDickinson FacsVantage+Diva Option auf CM-17 Agar mit 100 mg/l Cefotaxim, 100 mg/l Hygromycin. In diesem Fall wurde nur dort Pilzmycel gebildet, wo die Sporen gentechnisch verändert waren.

- 30 **Nachweis der genetischen Veränderung durch Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora***

Nachfolgend wird beispielhaft der Nachweis der Übertragung für pBinA-Hyg in *Blakeslea trispora* beschrieben. Der Nachweis der Übertragung der Abkömmlinge erfolgte analog.

- 5 200 ml MEP-Medium (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5) wurden mit 10^5 bis 10^7 Sporen der Transformante *Blakeslea trispora* GVO 3005 beimpft und 7 Tage bei 26 °C mit 200 Upm auf einem Rundschüttler inkubiert. Zum Nachweis der erfolgreichen Transformation wurde DNA aus dem Mycel isoliert (Peqlab Fungal DNA Mini Kit) und in einer PCR (Pro-
- 10 gramm: 94 °C 1 min, dann 30 Zyklen mit 1 min. 94°C, 1 min. 58 °C, 1 min. 72 °C) eingesetzt.

- Zum Nachweis des Hygromycinresistenzgens (hph) wurden die Primer hph-forward (5'-CGATGTAGGAGGGCGTGGATA, SEQ ID NO: 5) und
- 15 hph-reverse (5'-GCTTCTGCGGGCGATTTGTGT, SEQ ID NO: 6) verwendet. Das erwartete Fragment von hph wies eine Länge von 800 bp auf.

- Zur Amplifikation des Kanamycinresistenzgens nptIII und damit als Kontrolle auf Agrobakterien wurden die Primer nptIII-forward (5'-
- 20 TGAGAATATCACCGGAATTG, SEQ ID NO: 7) und nptIII-reverse (5'-AGCTCGACATACTGTTCTTCC, SEQ ID NO: 8) verwendet. Das erwartete Fragment von nptIII wies eine Länge von 700 bp auf.

- Zur Amplifikation eines Fragmentes des Glycerinaldehyd-3-phosphatdehydrogenasegens gpd1 und damit als Kontrolle auf *Blakeslea trispora* wurden die Primer MAT292 (5'-
- 25 GTGAATGGAAATCCCATCGCTGTC, SEQ ID NO: 9) und MAT293 (5'-AGTGGGTACTCTAAAGGCCATACC, SEQ ID NO: 10) verwendet. Das erwartete Fragment von gpd1 wies eine Länge von 500 bp auf.

Das Ergebnis der PCR der *Blakeslea trispora* DNA ist in Fig. 3 anhand eines Standard-Gels gezeigt. Die Spuren des Gels wurden folgendermaßen belegt:

- | | | |
|----|----------------------------------------|--------------------------------|
| 5 | 1) 100 bp Größenmarker (100 bp - 1 kb) | |
| | 2) B.t. GVO 3005 | primer nptIII-for / nptIII-rev |
| | 3) B.t. GVO 3005 | primer hph-for / hph-rev |
| | 4) B.t. GVO 3005 | primer MAT292 / MAT293 (gpd) |
| | 5) A.t. mit Plasmid pBinAHyg | primer nptIII-for / nptIII-rev |
| 10 | 6) A.t. mit Plasmid pBinAHyg | primer hph-for / hph-rev |
| | 7) B.t. 14272 WT | primer nptIII-for / nptIII-rev |
| | 8) B.t. 14272 WT | primer hph-for / hph-rev |
| | 9) B.t. 14272 WT | primer MAT292 / MAT293 (gpd) |
- 15 In der DNA von *Blakeslea trispora* wurde das Hygromycinresistenzgens (hph) und als Positivkontrolle Glycerinaldehyd-3-phosphatdehydrogenasegen (gpd1) nachgewiesen. nptIII konnte demgegenüber nicht nachgewiesen werden.
- 20 Somit wurde die genetische Veränderung von *Blakeslea trispora* durch *Agrobacterium*-vermittelte Transformation nachgewiesen.

Isolierung homokaryotischer GVO von *Blakeslea trispora*:

25 Herstellung homonukleater Stämme

Durch erfolgreichen Transfer des Vectors pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora* entstanden entstehen genetisch veränderte Organismen. In GVO von *Blakeslea trispora*. Jedoch liegen in *Blakeslea* liegen in allen Stadien des vegetativen und des sexuellen Zellzyklus mehrkernige Zellen vor. Daher erfolgte erfolgt die Insertion der VectorFremd-DNA in der Regel nur in einem Kern. Ziel ist es aber, dass,

30

Stämme von *Blakeslea* zu erhalten, bei denen die Insertion der Vector-Fremd-DNA in allen Kernen vorliegt., d.h. Ziel ist ein homonukleates rekombinantes Pilzmycel.

- 5 Zur Herstellung solcher homokaryotischer Zellen wurden zunächst Sporensuspensionen der rekombinanten Stämme mit MNNG behandelt. Hierfür wurde eine Sporensuspension mit 1×10^7 Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 µg/ml zugegeben. Die Dauer der Inkubation mit MNNG
10 wurde so gewählt, dass die Überlebensrate der Sporen ~5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und plattiert.

1) Herstellung homonukleater rekombinanter Stämme durch FACS (fluorescence-activated cell sorting)
15

- Ein geringer Anteil der Sporen von *Blakeslea trispora* bzw. der gentechnisch veränderten Stämme von *Blakeslea trispora* ist von Natur aus einkernig. Zur Herstellung homonukleater rekombinanter Stämme, die Fremd-DNA von pBinAHyg oder pBinAHyg-Abkömmlingen enthielten, wurden die
20 einkernigen Sporen durch FACS aussortiert und auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin plattiert. Die hier gebildeten Mycelien waren homonukleat. Zur Sortierung mit FACS wurden die Sporen eines 3 Tage alten Ausstriches mit 10 ml Tris-HCl 50mMol + 0,1% Span20 pro Agar-Platte abge-
25 schwemmt. Die Sporenkonzentration betrug 0,5 bis $0,8 \times 10^7$ Sporen pro ml. Zu 9 ml Sporensuspension wurden 1ml DMSO und 10 µl Syto 11 (Farbstoff-Stammlösung in DMSO Molecular Probes Nr.S-7573) zugegeben. Danach wurde 2 h bei 30°C gefärbt. Die Selektion und Ablage erfolgte mittels eines BectonDickinson FacsVantage+Diva Option. Die Selektion
30 erfolgt zuerst nach Größe, um einzelne Sporen von Aggregaten und Verunreinigungen zu trennen. Dann wurden diese Sporen nach ihrer Fluores-

zenz (Anregung = 488nm Emission = 530 nm) sortiert abgelegt. Die linke Schulter der Gauß-Kurve der Fluoreszenzhäufigkeitsverteilung enthielt die einkernigen Sporen.

- 5 Anschließend wurden die Sporen auf MEP-Agarplatten ausplattiert und neue Sporen erzeugt.

10 Diese Sporen wurden analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin plattiert, das zusätzlich Hygromycin enthielt.

Hierdurch wurden homokaryonte Zellen des Genotyps hyg^R und dar^- selektiert.

15

2) Herstellung homonukleater Stämme durch Kernreduktion und Selektion mit FACS

20 Zur Reduzierung der Anzahl von Kernen pro Spore wurde vor der Selektion eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt, und so durch chemische Mutagenese eine Kernreduktion erzielt.

25 Hierfür wurde zunächst eine Sporensuspension mit 1×10^7 Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 µg/ml zugegeben. Die Zeit der Inkubation in MNNG wurde so gewählt, dass die Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und nach der unter

1) beschriebenen Methode sortiert bzw. selektiert.

30

Alternativ konnten zur Reduktion der Kernzahl in den Sporen auch Röntgen – und UV-Strahlen eingesetzt werden, wie es von Cerdá-Olmedo und Patricia Reau in Mutation Res., 9 (1970), 369-384 beschrieben wurde.

5

3) Herstellung homonukleater Stämme durch Selektion auf rezessive Selektionsmarker

Als rezessiver Selektionsmarker zur Selektion homonukleater Mycelien kommt beispielsweise der rezessive Selektionsmarker *pyrG* in Frage.

10 Wildtyp-Stämme von *Blakeslea trispora* sind *pyrG*⁺. Diese Stämme können nicht in Gegenwart des Pyrimidin-Analogs 5-Fluororotat (FOA) wachsen, weil sie FOA durch die Orotidin-5'-monophosphatdecarboxylase zu lethalen Metaboliten umsetzen. Gentechnisch veränderte *Blakeslea*, die homonukleat *pyrG*⁻ sind, fehlt die Enzymaktivität Orotidin-5'-

15 monophosphatdecarboxylase. Folglich können diese *pyrG*⁻-Stämme 5-Fluororotat nicht verwerten. Die Stämme wachsen daher in Gegenwart von FOA und Uracil. Im Fall der Kopplung der Mutation *pyrG*⁻ und der Insertion von Fremd-DNA auf dem Kern einer einkernigen Spore, kann aus dieser Spore homonukleates rekombinantes Pilzmycel gebildet werden.

20

Zunächst wurde durch Insertion eines Fragments von *pyrG* (SEQ ID NO: 65) aus *Blakeslea trispora* in pBinAHyg das Plasmid pBinAHygBT*pyrG*-SCO (SEQ ID NO: 36, Fig. 4) erzeugt. Dieses Plasmid wurde in *Blakeslea trispora* transformiert und führte dort durch homologe Rekombination zur

25 Disruption von *pyrG*.

Homonukleate GVO von *Blakeslea trispora* mit dem Phänotyp *pyrG*⁻ wurden folgendermaßen selektiert. Zur Agrobakterium-vermittelten Transformation von pBinAHygBT*pyrG*-SCO wurde wie oben beschrieben auf MEP

30 (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin plattiert. Die Sporen der Transformanten

wurden mit 10 ml Tris-HCl 50mM + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug 0,5 bis $0,8 \times 10^7$ Sporen pro ml. Die Sporen wurden anschließend auf FOA-Medium mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin ausplattiert. FOA-Medium enthielt pro Liter 20 g Glucose, 1 g FOA, 50 mg Uracil, 200 ml Citrat-Puffer (0,5 M, pH 4,5) und 40 ml Spurensalzlösung nach Sutter, 1975, PNAS, 72:127). Homonukleate pyrG⁻-Mutanten zeigten Wachstum auf dem Uracil-haltigen FOA-Medium; aber kein Wachstum bei Plattierung auf FOA-Medium ohne Uracil. Auf die gleiche Weise wurden aus den im folgenden beschriebenen GVO von *Blakeslea trispora* zur Herstellung von Xanthophyllen homonukleate GVO hergestellt.

Alternativ ist es möglich die Sporen analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin zu plattieren, das zusätzlich Hygromycin enthält (Roncero et al., 1984, Mutation Research, 125: 195 - 204). Hierdurch werden homokaryonte Zellen des Genotyps hyg^R und dar⁻ selektiert. Nach diesem Prinzip werden homokaryonte Stämme von *Blakeslea trispora* mit dem Phänotyp hyg^R und dar⁻ erzeugt.

20

Ausführungsbeispiele zur Herstellung von gentechnisch veränderten Organismen von *Blakeslea trispora* für die Herstellung von Carotinoiden und Carotinoidvorstufen

Die Erzeugung der im folgenden genannten Plasmide erfolgte durch die Methode „overlap-extension PCR“ und durch anschließende Insertion der Amplifikationsprodukte in das Plasmid pBinAHyg. Die Methode „overlap-extension PCR“ erfolgte wie in Innis et al. (Eds.) PCR protocols: a guide to methods and applications, Academic Press, San Diego beschrieben. Die Transformation der pBinAHyg-Abkömmlinge und die Herstellung homonukleater gentechnisch veränderter Stämme von *Blakeslea trispora* erfolgte wie oben beschrieben.

Gentechnisch veränderte Stämme von *Blakeslea trispora* zur Herstellung von Zeaxanthin

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Zeaxanthin verwendet, codieren also u.a. Hydroxylasen (crtZ):

- 5 - p-tef1-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ (SEQ ID NO: 70) aus *Haematococcus pluvialis* Flotow NIES-144 (Accession No. AF162276) unter Kontrolle des ptef1 Promotors aus *Blakeslea trispora* (Seq. pBinAHygBTpTEF1-HPcrtZ, SEQ ID NO: 37, Fig. 5);
- 10 - p-carRA-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHyg-BTpcarRA-HPcrtZ, SEQ ID NO: 38, Fig. 6)
- 15 - p-carB-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTpcarB-HPcrtZ, SEQ ID NO: 39, Fig. 7)
- 20 - p-carRA-HPcrtZ-TAG-3'carA-IR, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora*. Stromabwärts des Gens der Hydroxylase ist eine Inverted-Repeat-Struktur lokalisiert, die aus dem 3'-Ende von carA und der stromabwärts von carA gelegenen Region stammt (IR, SEQ ID NO: 74, 'Inverted Repeat 1'
- 25 ca. 350 bp von carA, dann ca. 200 bp 'Loop' und anschließend ca. 350 bp 'Inverted Repeat 2') (Seq. pBinAHyg-BTpcarRA-HPcrtZ-TAG-3'carA-IR, SEQ ID NO: 40, Fig. 8);
- 30 - p-carRA-HPcrtZ-GCG-3'carA-IR, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora*. Das Gen der

- Hydroxylase ist mit einer Inverted-Repeat-Struktur fusioniert, die aus dem 3'-Ende von *carA* und der stromabwärts von *carA* gelegenen Region stammt (IR, SEQ ID NO: 74, 'Inverted Repeat 1' ca. 350 bp von *carA*, dann ca. 200 bp 'Loop' und anschließend ca. 350 bp 'Inverted Repeat 2'). Das abgeleitete Fusionsprotein besteht folglich aus der Hydroxylase von *Haematococcus pluvialis* und dem Carboxyterminus von *CarA* aus *Blakeslea trispora* (Seq. pBinAHyg-BTpcarRA-HPcrtZ-GCG-3'carA-IR, SEQ ID NO: 41, Fig. 9);
- p-tef1-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ (SEQ ID NO: 71) aus *Erwinia uredovae* 20D3 (Accession No. D90087) unter Kontrolle des ptef1 Promotors (Seq. pBinAHygBTpTEF1-EUcrtZ, SEQ ID NO: 42, Fig. 10);
 - p-carRA-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ aus *Erwinia uredovae* 20D3 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-EUcrtZ, SEQ ID NO: 43, Fig. 11);
 - p-carB-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ aus *Erwinia uredovae* 20D3 unter Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTpcarB-EUcrtZ, SEQ ID NO: 44, Fig. 12);
 - p-gpdA-HPcrtZ-t-crtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des gpdA Promotors und des Terminators t-crtZ; d.h. des stromabwärts von crtZ aus *Haematococcus pluvialis* Flotow NIES-144 gelegenen Sequenzabschnitts (SEQ ID NO: 73) (Seq. pBinAHyg-gpdA-HPcrtZ-tcrtZ, SEQ ID NO: 45, Fig. 13).
 - p-gpdA-BTcarR-HPcrtZ-BTcarA, enthaltend Genfusion aus Genen der Lycopincyclase carR aus *Blakeslea trispora*, der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 und der

Phytoensynthase *carA* aus *Blakeslea trispora* unter Kontrolle des *gpdA* Promotors aus *Aspergillus nidulans* (Seq. *pBinAHyg-carR_crtZ_carA*, SEQ ID NO: 46, Fig. 14);

5 Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Canthaxanthin

Folgende Plasmide (Abkömmlinge von *pBinAHyg*) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Canthaxanthin verwendet, codieren also u.a. Ketolasen (*crtW*):

- 10 - *p-tef1-NPcrtW*, enthaltend das Gen der Ketolase *NPcrtW* (SEQ ID NO: 72) aus *Nostoc punctiforme* PCC73102 (ORF148, Accession No. NZ_AABC01000196) unter Kontrolle des *ptef1* Promotors aus *Blakeslea trispora* (Seq. *pBinAHygBTpTEF1-NpucrtW*, SEQ ID NO: 47, Fig. 15);
- 15 - *p-carRA-NPcrtW*, enthaltend das Gen der Ketolase *NPcrtW* aus *Nostoc punctiforme* PCC73102 unter der Kontrolle des Promotors *pcarRA* aus *Blakeslea trispora* (Seq. *pBinAHygBTpcarRA-NpucrtW*, SEQ ID NO: 48, Fig. 16);
- 20 - *p-carB-NPcrtW*, enthaltend das Gen der Ketolase *NPcrtW* aus *Nostoc punctiforme* PCC73102 unter der Kontrolle des Promotors *pcarB* aus *Blakeslea trispora* (Seq. *pBinAHygBTpcarB-NpucrtW*, SEQ ID NO: 49, Fig. 17);

25 Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Astaxanthin

Folgende Plasmide (Abkömmlinge von *pBinAHyg*) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Astaxanthin verwendet, codieren also u.a. für Hydroxylasen (*crtZ*) und Ketolasen (*crtW*):

- p-carRA-HPcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 und das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 (ORF148, Accession No. NZ_AABC01000196) beide jeweils unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-HPcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 50, Fig. 18);
- p-carRA-EUcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase EUcrtZ aus *Erwinia uredovae* 20D3 (Accession No. D90087) und das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 beide jeweils unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 51, Fig. 19);

15 Klonierung und Sequenzanalyse von Genen und Promotoren, die beispielhaft für die gentechnische Veränderung von *Blakeslea trispora* genutzt werden können.

Nachfolgend werden beispielhaft die Klonierung und Sequenzierung verschiedener Gene und Promotoren aus *Blakeslea trispora* beschrieben.

20

Klonierung und Sequenzanalyse ptef1

Die Klonierung von p-tef aus *Blakeslea trispora* erfolgte auf der Grundlage einer bereits in GenBank veröffentlichten Sequenz des Strukturgens für den Translations-Elongationsfaktor 1- α aus *Blakeslea trispora* (AF157235). Ausgehend von dem Sequenzeintrag AF157235 wurden Primer für die inverse PCR ausgewählt, um die stromaufwärts des Strukturgens gelegene Promotoregion zu amplifizieren und zu sequenzieren.

In der inversen nested PCR an 200 ng XhoI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272 wurde ein 3000-bp-Fragment in folgendem Ansatz erhalten: Matrizen-DNA (1 μ g genomi-

30

sche DNA von *Blakeslea trispora* ATCC 14272) Primer MAT344 5'-GGCGTACTTGAAGGAACCCTTACCG-3' (SEQ ID NO: 63) und MAT 345 5'-ATTGATGCTCCCGGTCACCGTGATT-3' (SEQ ID NO: 64) je 0,25 µM, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zu-
5 gabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus). Der Sequenzabschnitt, der stromaufwärts des vermutlichen Startcodons des Gens *tef1* innerhalb 3000-bp-Fragmentes liegt, wurde als Promotor *ptef1* bezeichnet.

10

Klonierung Sequenzanalyse des Gens der HMG-CoA-Reduktase aus *Blakeslea trispora*

Zunächst wurde mit dem Cosmidvektor pANsCos1 eine Genbank von *Blakeslea trispora* ATCC 14272, Mating Type (–) hergestellt. Der Vektor wurde durch Spaltung mit XbaI linearisiert und anschließend dephosphoryliert.
15 Eine weitere Spaltung mit BamHI schuf die Insertionsstelle, in welche die mit Sau3AI partiell gespaltene und dephosphorylierte genomische DNA von *Blakeslea trispora* ligiert wurde. Die derart gebildeten Cosmide wurden anschließend *in vitro* verpackt und in *Escherichia coli* übertragen.

20 Auf der Grundlage der bekannten Sequenz eines Fragmentes des HMG-CoA-Reduktase codierenden Gens aus *Blakeslea trispora* (Eur. J. Biochem 220, 403–408 (1994)) wurde eine 315-bp-DNA-Sonde durch folgende PCR hergestellt. Reaktionsansatz: 1 µg genomische DNA von *Blakeslea trispora* ATCC 14272, Primer MAT314 5'-
25 CCGATGGCGACGACGGAAGGTTGTT-3' [SEQ ID NO 79] und MAT315 5'-CATGTTCATGCCCATTCATCACCT-3' [SEQ ID NO 80] je 0,25 µM, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 58 °C, 30 s. 72 °C, 30 s, 95 °C, 30 s (30
30 Zyklen); 72 °C, 10 min (1 Zyklus).

Mit dieser DNA-Sonde wurde die Cosmid-Genbank durchmustert. Es wurde ein Klon identifiziert, dessen Cosmid mit der DNA-Sonde hybridisierte. Die Insertion dieses Cosmids wurde sequenziert. Die DNA-Sequenz enthielt einen Abschnitt, der dem Gen einer HMG-CoA-Reduktase zugeordnet wurde [HMG-CoA-Red.gb].

Klonierung und Sequenzanalyse *carB*

(*carB* = Gen der Phytoendesaturase aus *Blakeslea trispora*)

Aus dem Sequenzvergleich der Peptidsequenzen von Phytoendesaturasen und dem Vergleich der zugehörigen DNA-Sequenzen von *Phycomyces blakesleeanus*, *Cercospora nicotianae*, *Phaffia rhodozyma* und *Neurospora crassa* wurden die degenerierten Primer MAT182 5'-GCNGARGGNATHHTGGTA-3' (SEQ ID 52) und MAT192 5'-TCNGCNAGRAADATRTTGTG-3' (SEQ ID 53) abgeleitet. Die PCR wurde in 100 µl Ansätzen durchgeführt. Diese enthielten 200 ng genomische DNA von *Blakeslea trispora* ATCC14272, 1 µM MAT182, 1 µM MAT192, 100 µM dNTP, 10 µl Pfu-Polymerasepuffer 10x, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C), H₂O ad 100 µl.

Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 40 °C, 30 s, 72 °C, 30 s, 95 °C, 30 s (35 Zyklen); 72 °C, 10 min (1 Zyklus).

Hiermit wurde ein 358-bp-Fragment erhalten, dessen abgeleitete Peptidsequenz Ähnlichkeit zu den Sequenzen der Phytoendesaturasen aufwies. Durch die Methode der inversen PCR (Innis et al. in PCR protocols: a guide to methods and applications. 1990. S. 219-227) wurden nach dem Prinzip des Chromosome-Walking die Genregionen stromaufwärts und stromabwärts des 350-bp-Fragmentes folgendermaßen amplifiziert, kloniert und sequenziert:

(i) ein 1,1-kbp-Fragment durch PCR mit den Primern MAT219 5'-AAGTGACACCGGTTACACGCTTGTCTT-3' (SEQ ID 54) und MAT

- 220 5'-GCTTATCACCATCTGTTACCTCCTTGC-3' (SEQ ID 55) erhalten aus 200 ng EcoRI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272, 0,25 µM MAT219, 0,25 µM MAT220, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus),
- 5
- (ii) ein 2,9-kbp-Fragment durch PCR mit den Primern MAT219 und MAT220 erhalten aus 200 ng XbaI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272, 0,25 µM MAT219, 0,25 µM MAT220, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s, 72 °C, 3 min, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus);
- 10
- 15

Der klonierte Sequenzabschnitt ist schematisch in Fig. 20 (SEQ ID NO 77) dargestellt. Die Sequenzierung erfolgte in Strang- und Gegenstrangrichtung mit den klonierten Fragmenten sowie mit den PCR-Produkten. Die Sequenz des klonierten Sequenzabschnitts ist in Fig. 21 (SEQ ID NO 78) gezeigt.

20

Sequenzvergleiche

Die Nukleotidsequenz von carB und die Peptidsequenz des abgeleiteten Proteins CarB wurden mit den bekannten Sequenzen verwandter Proteine verglichen. Zum Sequenzvergleich wurden die Programme GAP und BESTFIT eingesetzt.

25

30 CarB - Identische Aminoacylreste nach GAP

Programmeinstellungen:

Gap Weight: 8

Length Weight: 2

Average Match: 2.912

Average Mismatch: -2.003

- 5 Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 72,491

Phaffia rhodozyma: 50,460

Neurospora crassa: 47,943

- 10 *Cercospora nicotianae*: 47,740

CarB -Identische Aminoacylreste nach BESTFIT

Programmeinstellungen:

Gap Weight: 8

- 15 Length Weight: 2

Average Match: 2.912

Average Mismatch: -2.003

Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

- 20 *Phycomyces blakesleeanus*: 73,380

Phaffia rhodozyma: 53,175

Neurospora crassa: 51,896

Cercospora nicotianae: 50,791

- 25 **carB - Identische Basen nach GAP**

Programmeinstellungen:

Gap Weight: 50

Length Weight: 3

Average Match: 10.000

- 30 Average Mismatch: 0.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 64,853

Cercospora nicotianae: 50,143

5 *Phaffia rhodozyma*: 43,179

Neurospora crassa: 42,130

carB -Identische Basen nach BESTFIT

Programmeinstellungen:

10 Gap Weight: 50

Length Weight: 3

Average Match: 10.000

Average Mismatch: -9.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

15

Phycomyces blakesleeanus: 68,926

Phaffia rhodozyma: 62,403

Neurospora crassa: 60,230

Cercospora nicotianae: 56,884

20

Klonierung zur Expression von carB

Zur Klonierung und Expression von carB aus *Blakeslea trispora* wurden von dem oben beschriebenen klonierten Sequenzabschnitt aus *Blakeslea trispora* in sechs Leserastern die möglichen Proteinsequenzen abgeleitet.

25 Diese Proteinsequenzen wurden mit den Sequenzen der Phytoendesaturasen aus *Phycomyces blakesleeanus*, *Phaffia rhodozyma*, *Neurospora crassa*, *Cercospora nicotianae* verglichen. Auf der Grundlage des Sequenzvergleiches wurden im klonierten Sequenzabschnitt der genomischen DNA von *Blakeslea trispora* drei Exons identifiziert, die zusammen-

30 gefügt eine codierende Region ergeben, deren abgeleitetes Genprodukt über die gesamte Länge 72,7% identische Aminoacylreste mit der Phy-

toendesaturase CarB aus *Phycomyces blakesleeanus* aufweist. Dieser Sequenzabschnitt aus drei möglichen Exons und zwei möglichen Introns wurde daher als Gen *carB* bezeichnet. Zur Überprüfung der vorhergesagten Genstruktur wurde die codierende Sequenz von *carB* aus *Blakeslea trispora* durch PCR mit cDNA von *Blakeslea trispora* als Matrize und mit den Primern Bol1425 5'-AGAGAGGGATCCTTAAATGCGAATATCGTTGC-3' (SEQ ID 56) und Bol1426 5'-AGAGAGGGATCCATGTCTGATCAAAAGAAGCA-3' (SEQ ID 57) erzeugt. Das erhaltene DNA-Fragment wurde sequenziert. Die Lokalisation von Exons und Introns wurde durch Vergleich der cDNA mit der genomischen DNA von *carB* bestätigt. In Fig. 21 ist die codierende Sequenz von *carB* schematisch dargestellt. Zur Expression von *carB* in *Escherichia coli* wurde zunächst die NdeI-Schnittstelle in *carB* durch die Methode overlap extension PCR entfernt sowie am 5'-Ende des Gens eine NdeI-Schnittstelle und am 3'-Ende eine BamHI-Schnittstelle eingefügt. Das erhaltene DNA-Fragment wurde mit dem Vektor pJOE2702 ligiert. Das erhaltene Plasmid wurde als pBT4 bezeichnet und zusammen mit pCAR-AE in *Escherichia coli* XL1-Blue kloniert. Die Expression erfolgte durch Induktion mit Rhamnose. Der Nachweis der Enzymaktivität erfolgte durch Nachweis der Lycopinsynthese via HPLC. Die Klonierungsschritte sind im folgenden beschrieben:

PCR 1.1:

Ca. 0,5 µg cDNA von *Blakeslea trispora*, 0,25 µM MAT350 5'-ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' (SEQ ID 58), 0,25 µM MAT244 5'-GTTCCAATTGGCCACATGAAGAGTAAGACAGGAAACAG-3' (SEQ ID 59), 100 µM dNTP, 10 µl Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100µL.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30 s, 9. 72 °C 10min
 Zyklen: (1-2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

5 **PCR1.2:**

Ca. 0,5 µg cDNA von *Blakeslea trispora*, 0,25 µM MAT243 5'-CCTGTCTTACTCTTCATGTGGCCAATTGGAACCAACAC-3' (SEQ ID 60), 0,25 µM MAT353 5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' (SEQ ID 61),
 10 100 µM dNTP, 10 µl Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100 µL.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30s, 9. 72 °C 10min
 15 Zyklen: (1 -2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

Reinigung der PCR-Fragmente aus PCR 1.1, 1.2

Dazu wurde PCR 2 zur Herstellung der codierenden Sequenz von *carB* aus *Blakeslea trispora* für die Klonierung in pJOE2702 durchgeführt:

20 Ca. 50 ng Produkt aus PCR 1.1 und ca. 50 ng Produkt aus PCR1.2 mit 0,25 µM MAT350 5'-ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' (SEQ ID NO 58), 0,25 µM MAT353 5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' (SEQ ID NO 61),
 25 100 µM dNTP, 10 µL Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100 µL.

Temperaturprofil:

1. 95°C 10 min, 2. 85 °C 5 min, 3. 59 °C 30 s, 4. 72 °C 2 min, 5. 95 °C 30 s, 6. 72°C 10 min
 30 Zyklen: (1-2.) 1x, (3-5.) 22x, (6.) 1x

Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Ligation in Vektor pPCR-Script-Amp, eine Klonierung in *Escherichia coli* XL1-Blue, Sequenzierung der Insertion, Spaltung mit NdeI und BamHI sowie eine Ligation in pJOE2702. Das erhaltene Plasmid wurde als pBT4 bezeichnet.

Charakterisierung und Nachweis der Enzymaktivität von CarB (Phytoendesaturase)

Das von carB abgeleitete Genprodukt wurde als CarB bezeichnet. CarB weist auf Grundlage der Peptidsequenzanalyse folgende Eigenschaften auf:

	Länge:	582 Aminoacylreste
	Molekulare Masse:	66470
	Isoelektrische Punkt:	6,7
15	Katalytische Aktivität:	Phytoendesaturase
	Edukt:	Phytoen
	Produkt:	Lycopin
	EC-Nummer:	EC 1.14.99-

Der Nachweis der Enzymaktivität erfolgte in vivo. Wenn das Plasmid (pCAR-AE) in *Escherichia coli* XL1-Blue übertragen wird, entsteht der Stamm *Escherichia coli* XL1-Blue (pCAR-AE). Dieser Stamm synthetisiert Phytoen. Wenn zusätzlich das Plasmid pBT4 in *Escherichia coli* XL1-Blue übertragen wird, entsteht der Stamm *Escherichia coli* XL1-Blue (pCAR-AE)(pBT4). Da ausgehend von carB eine enzymatisch aktive Phytoendesaturase gebildet wird, produziert dieser Stamm Lycopin.

Die Plasmide pCAR-AE und pBT4 wurden daher in *Escherichia coli* übertragen. Nach Wachstum in Flüssigkultur wurden die Carotinoide aus den Zellen extrahiert und charakterisiert (vgl. oben).

Durch HPLC Analyse wurde nachgewiesen, daß der Stamm *Escherichia coli* XL1-Blue (pCAR-AE) Phytoen und der Stamm *Escherichia coli* XL1-Blue (pCAR-AE)(pBT4) Lycopin produziert. CarB weist folglich die Enzymaktivität einer Phytoendesaturase auf.

5

Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Phytoen

Nachfolgend werden beispielhaft die Herstellung von gentechnisch veränderten Organismen zur Herstellung von Phytoen beschrieben.

10

Vector pBinAHyg Δ carB zur Erzeugung von carB⁻-Mutanten von *Blakeslea trispora*

Für die Deletion von carB in *Blakeslea trispora* wurde der Vektor pBinAHyg Δ carB (SEQ. ID. NO:62, Fig. 22) konstruiert. Der Vorläufer von pBinAHyg Δ carB ist pBinAHyg (SEQ. ID. NO:3, Fig. 2). pBinAHyg wurde folgendermaßen konstruiert:

Aus dem Plasmid pANsCos1 (SEQ. ID. NO:4, Fig. 1, Osiewacz, 1994, Curr. Genet. 26:87-90) wurde die gpdA-hph Kasette als BglII/HindIII Fragment isoliert und in das BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet und enthält das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors und des trpC Terminators aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind.

25

Die Amplifikation der codierenden Sequenz von carB mit den Primern MAT350 (SEQ ID NO 58) und MAT353 (SEQ ID NO 61) mittels PCR wurde mit den folgenden Parametern durchgeführt:

30 50 ng pBT4 mit 0,25 μ M MAT350 5'-ACTTTATTGGATCCTTAAAT-GCGAATATCGTTGCTGC-3', 0,25 μ M MAT353 5'-

CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3', 100 μ M dNTP, 10 μ L Pfu-Polymerase-Puffer, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 μ L H₂O

Temperaturprofil:

- 5 1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 30s, 4. 72°C 2 min, 5. 95 °C 30s, 6. 72 °C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

- 10 Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Spaltung mit HindIII, eine weitere Reinigung des 364-bp-HindIII-Fragments-carB, gefolgt von einer Spaltung von pBinAHyg mit HindIII, eine Ligation von 364-bp-HindIII-Fragments-carB in pBinAHyg, eine Transformation des Vektors in Escherichia coli und eine Isolierung des Konstruktes und Bezeichnung als pBinAHyg Δ carB wie oben beschrieben. Alternativ
- 15 erfolgte eine partielle Spaltung mit HindIII und die Klonierung eines größeren HindIII-Fragmentes aus carB in pBinAHyg zur Herstellung von pBinAHyg Δ carB.

20 Erzeugung von carB⁻-Mutanten von *Blakeslea trispora*

- Zunächst wurde das Plasmid pBinAHyg Δ carB in den Agrobakterienstamm LBA 4404 übertragen, z. B. durch Elektroporation (vgl. oben). Anschließend wurde das Plasmid von *Agrobacterium tumefaciens* LBA 4404 in *Blakeslea trispora* ATCC 14272 und in *Blakeslea trispora* ATCC 14271
- 25 übertragen (vgl. oben). Der erfolgreiche Nachweis des Gentransfers in *Blakeslea trispora* erfolgte über Polymerase-Kettenreaktion nach folgendem Protokoll:

- Ca. 0,5 μ g DNA aus *Blakeslea trispora* ATCC 14272 carB⁻ bzw. ATCC 14271 carB⁻ wurden mit 0,25 μ M Primer hph forward 5'-
- 30 CGATGTAGGAGGGCGTGGATA-3' (SEQ ID NO 5), 0,25 μ M Primer hph reverse 5'-GCTTCTGCGGGCGATTTGTGT-3' (SEQ ID NO 6), 100 μ M

dNTP, 10 µL Herculase-Polymerase-Puffer, 2,5 U Herculase-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µl H₂O umgesetzt.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6. 72 °C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

Als Negativkontrolle wurde eine Amplifikation des Kanamycinresistenzgens aus *Agrobacterium* versucht. Dazu wurden folgende PCR-Bedingungen verwendet:

Ca. 0,5 µg DNA aus *Blakeslea trispora* ATCC 14272 *carB*⁻ bzw. ATCC 14271 *carB*⁻ wurden mit 0,25 µM Primer nptIII forward 5'-TGAGAATATCACCGGAATTG-3' (SEQ ID NO 7), 0,25 µM Primer nptIII reverse 5'-AGCTCGACATACTGTTCTTCC-3' (SEQ ID NO 8), 100 µM dNTP, 10 µL Herculase-Polymerase-Puffer, 2,5 U Herculase-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µL H₂O umgesetzt.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6. 72 °C 10 min-

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

C) Produktion von Carotinoiden und Carotinoidvorstufen mit *Blakeslea trispora*

Zur Produktion der Carotinoide Zeaxanthin, Canthaxanthin, Astaxanthin und Phytoen wurden die entsprechenden gentechnisch veränderten *Blakeslea trispora* (+) und (-) Stämme fermentiert, das produzierte Carotinoid mittels HPLC Analyse nachgewiesen und isoliert.

Das Flüssigmedium zur Produktion von Carotinoiden enthielt pro Liter: 19 g Maismehl, 44 g Sojamehl, 0,55 g KH_2PO_4 , 0,002 g Thiaminhydrochlorid, 10 % Sonnenblumenöl. Der pH wurde mit KOH auf 7,5 eingestellt.

- 5 Zur Herstellung der Carotinoiden wurden Schüttelkolben mit Sporensuspensionen von (+) und (-) Stämmen der GVO von *Blakeslea trispora* beimpft. Die Schüttelkolben wurden bei 26 °C mit 250 rpm für 7 Tage inkubiert. Alternativ wurde zu Mischungen der Stämme nach 4 Tagen Trispor-säuren zugegeben und weitere 3 Tage inkubiert. Die Endkonzentration der
- 10 Trispor-säuren betrug 300 - 400 µg/ml.

Extraktion und Analytik

Extraktion:

1. Entnahme von 10 ml Kultursuspension
- 15 2. Zentrifugation, 10 min, 5.000 x g
3. Verwerfen des Überstandes
4. Resuspendierung des Pellets in 1 ml Tetrahydrofuran (THF) durch Vortexen
5. Zentrifugation, 5 min, 5.000 x g
- 20 6. Abnahme der THF-Phase
7. Wiederholung der Schritte 4.-6. (2 x)
8. Vereinigung der THF-Phasen
9. Zentrifugation der vereinigten THF-Phasen 5 min bei 20.000 x g, um Reste der wäßrigen Phase abzutrennen

25

Analytik

Messung von Phytoen mittels HPLC

- Säule: ZORBAX Eclipse XDB-C8, 5 µm, 150*4,6 mm
- 30 Temperatur: 40 °C
- Flußrate: 0,5 ml/min

Injektionsvolumen: 10 µl

Detektion: UV 220 nm

Stoppzeit: 12 min

Nachlaufzeit: 0 min

5 Maximaldruck: 350 bar

Eluent A: 50 mM NaH₂PO₄, pH 2,5 mit Perchlorsäure

Eluent B: Acetonitril

Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
10	0	50	50	0,5
	12	50	50	0,5

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden jeweils 50 - 1000 mg/l eingewogen und in THF gelöst. Als Standard wurde Phytoen verwendet, welches unter den gegebenen Bedingungen eine Retentionszeit von 7,7 min. aufweist.

20 **Messung von Lycopin, β-Carotin, Echinenon, Canthaxanthin, Cryptoxanthin, Zeaxanthin und Astaxanthin mittels HPLC**

Säule: Nucleosil 100-7 C18, 250*4,0 mm (Macherey & Nagel)

Temperatur: 25 °C

Flußrate: 1,3 ml/min

25 Injektionsvolumen: 10 µl

Detektion: 450 nm

Stoppzeit: 15min

Nachlaufzeit: 2 min

Maximaldruck: 250 bar

30 Eluent A: 10% Aceton, 90% H₂O

Eluent B: Aceton

Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
	0	30	70	1,3
	10	5	95	1,3
5	12	5	95	1,3
	13	30	70	1,3

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden jeweils 10 mg eingewogen und in 100 ml THF gelöst. Als Standard wurden folgende Carotinoide mit folgenden Retentionszeiten eingesetzt β-Carotin (12,5 min), Lycopin (11,7 min), Echinenon (10,9 min), Cryptoxanthin (10,5 min), Canthaxanthin (8,7 min), Zeaxanthin (7,6 min) und Astaxanthin (6,4 min) [s. Fig. 23].

Produktion von Zeaxanthin mit gentechnisch veränderten Stämmen von *Blakeslea trispora*

Nachfolgend wird beispielhaft die Herstellung von Zeaxanthin mit gentechnisch veränderten Organismen (GVO) von *Blakeslea trispora* beschrieben.

Durch Agrobakterium-vermittelte Transformation wurde der Vektor pBinA-HygBTpTEF1-HPcrtZ in *Blakeslea trispora* übertragen (s.o.). Ein Hygromycin-resistenter Klon wurde isoliert und auf eine Kartoffel-Glucose-Agarplatte (Merck KGaA, Darmstadt) übertragen.

Nach drei Tagen Inkubation bei 26°C wurde ausgehend von dieser Platte eine Sporensuspension hergestellt. Ein 250-ml-Erlenmeyerkolben ohne Schikanen mit 50 ml Growth-Medium (Maismehl 47 g/l, Sojamehl 23 g/l, KH₂PO₄ 0,5 g/l, Thiamin-HCl 2.0 mg/l, pH mit NaOH vor der Sterilisation auf 6,2–6,7 eingestellt) wurde mit 1x10⁵ Sporen beimpft. Diese Vorkultur inkubierte 48 Stunden bei 26 °C und 250 upm. Für die Hauptkultur

wurde ein 250-ml-Erlenmeyerkolben ohne Schikane enthaltend 40 ml Produktionsmedium mit 4 ml der Vorkultur beimpft und 8 Tage bei 26 °C und 150 upm inkubiert. Das Produktionsmedium enthielt Glucose 50 g/l, Casein Acid Hydrolysate 2 g/l, Hefeextrakt 1 g/l, L-Asparagin 2 g/l, KH₂PO₄ 1,5 g/l, MgSO₄ x 7 H₂O 0,5 g/l, Thiamin-HCl 5 mg/l, Span20 10 g/l, Tween 80 1 g/l, Linolsäure 20 g/l, Maisquellwasser 80 g/l. Nach 72 Stunden erfolgte die Zugabe von Kerosin in einer Endkonzentration von 40 g/l Kerosin.

Nach der Ernte der Kulturen werden die verbliebenen ungefähr 35 ml Kultur mit Wasser auf 40 ml aufgefüllt. Anschließend werden die Zellen im Hochdruckhomogenisator, Typ Micron Lab 40, Fa. APV Gaulin, 3 x bei 1500 bar aufgeschlossen.

Die Suspension mit den aufgeschlossenen Zellen wurde mit 35 ml THF versetzt und 60 min bei RT im Dunkeln bei 250 upm geschüttelt. Danach wurden 2 g NaCl zugegeben und das Gemisch nochmals geschüttelt. Der Extraktionsansatz wurde dann 10 min bei 5000 x g zentrifugiert. Die gefärbte THF-Phase wurde abgenommen, die Zellmasse war vollständig entfärbt.

Die THF-Phase wurde am Rotationsverdampfer bei 30 mbar und 30 °C auf 1 ml eingeeengt und danach nochmals in 1 ml THF aufgenommen. Nach Zentrifugation 5 min bei 20 000 x g wurde ein Aliquot der oberen Phase entnommen und durch HPLC analysiert (Fig. 24, Fig. 23).

D) Aufarbeitung und Isolierung der Carotinoide bzw. des Nahrungsmittels

Die oben unter A) angegebenen Kulturbrühen wurden wie nachfolgend aufgearbeitet, um hochreine Carotinoide und ein entsprechendes Nahrungsmittel zu erhalten.

Der Carotinoidgehalt der Kulturbrühen 1, 2, 3 betrug zwischen 0,5 und 1,5 g/L.

D1) Beispiel gemäß Variante a) IIA und Variante b) IIA bzw. IIB

- 5 Die Kulturen mit identischen Medien (insgesamt ca. 1 L) wurden am Ende des Kultivierungszeitraums vereinigt und mit Hilfe eines Dispergiergeräts (Ultra-Turrax ®) homogenisiert.

Die Feststoffkonzentration in den Medien 1 und 2 betrug 37 g/l bzw. 11
10 g/L. Die Entwässerung der Kulturbrühe erfolgte durch eine Zentrifuge. Bei hohen Zellkonzentrationen bzw. hohem Feststoffgehalt des Mediums kann die Kulturbrühe auch ohne vorherige fest-flüssig-Trennung weiterverarbeitet werden (Medium 3: 127 g Feststoff/L. Nach vorheriger Homogenisation mit einem Dispergiergerät (Ultra-Turrax ®) und unter ständigem Rühren
15 der Suspension wurde die Zellmasse über eine Schlauchpumpe auf den Trockner aufgegeben. Die Eindüsung in den Zylinder des Laborsprüh-trockners erfolgte dabei über eine Zweistoffdüse mit dem Durchmesser 2,0 mm. Eingedüst wurde mit 2 bar und 4,5 Nm³/h Stickstoff. Die Temperatur am Eintritt betrug ca. 125°C bis 127°C. Das Trocknungsgas war Stickstoff
20 mit einer Flussrate von 22 Nm³/h. Die Austrittstemperatur betrug zwischen 59°C und 61°C. Bei jeder der drei Fermentationsbrühen konnte am Zyklon des Sprühtrockners rieselfähiges Produkt abgeschieden werden. Die Wandbeläge im Turm (sofern vorhanden) platzten automatisch von der Gefäßwand ab und werden als unproblematisch eingestuft.

25

Es wurden zwischen 8 und 100 g pulveriges Nahrungsmittel erhalten, welches direkt als Tierfuttermittel verwendet werden könnte. Es enthielt ca. 1-10 % Carotinoide bezogen auf das Trockengewicht. Die Restfeuchte betrug weniger als 5%.

30

Beispiel gemäß Variante b) IIC**D2) Extraktion mit Tetrahydrofuran**

Die Zellen aus je 40 ml der Kulturbrühen 1, 2, 3 wurden 3 x bei 1500 bar
5 durch einen Hochdruckhomogenisator, Typ Micron Lab 40, Fa. APV Gau-
lin aufgeschlossen. Je 20 ml der Suspensionen mit den aufgeschlossenen
Zellen wurden mit 20 ml Tetrahydrofuran versetzt und 30 min, bei 30°C im
Rundschüttler bei 200 Upm geschüttelt. Danach wurden 2 g NaCl zuge-
10 setzt und zur Phasentrennung 5 min bei 5000 x g zentrifugiert. Die THF-
Phase wurde abgenommen. Danach wurde die wässrige Phase nochmals
mit 20 ml THF extrahiert. Die Extrakte wurde vereinigt. Die Carotinoidkon-
zentration wurde durch HPLC quantifiziert.

D3) Extraktion mit Methylenchlorid

15 Die Biomassenabtrennung aus der Kulturbrühe (200 mL) erfolgte durch
Zentrifugation bei 5.000 x g für 10 min. in einer Laborzentrifuge.

Die abgetrennte Biofeuchtmasse (jeweils ca. 10 g bis 100 g) wurde mit 10
- 100 mL Wasser vermischt, um wasserlösliche Komponenten zu entfer-
20 nen. Die Biomasse wurde abgetrennt (Laborzentrifuge) und danach mit
Dampf (T = 121, t = 30 min, 1 bar) im Autoklaven sterilisiert und so die
Zellen aufgeschlossen.

Zu den Zelltrümmern wurden 25 - 250 g Methylenchlorid zugegeben und
25 das Carotinoid aus der Biomasse mittels Ausschütteln extrahiert. Die Bio-
masse wurde in einer Laborzentrifuge abgetrennt.

Es wurde ein Lösungsmitteltausch von Methylenchlorid zu Methanol
durchgeführt, wozu die Carotinoidlösung ca. vier Stunden bei 40°C bis
30 60°C gehalten und über diesen Zeitraum kontinuierlich mit insgesamt 20 -
200 mL Methanol versetzt wurde. Methylenchlorid wurde dabei als Lö-

sungsmittel zurückgewonnen. Erste Carotinoid Kristalle fielen aus. Anschließend wurde langsam, über 6 h auf ca. 10 °C abgekühlt, wobei die Carotinoid Kristalle an Größe und Anzahl zunahmen. . Danach wurde die Mutterlauge abfiltriert und die Carotinoid Kristalle getrocknet. Ein Teil der
5 Mutterlauge kann zum Lösungsmitteltausch wiederverwendet werden. Der andere Teil wird destilliert und das so gereinigte Methanol im Lösungsmitteltausch wiederverwendet.

Es wurden 0,0,08 g bis 0,24 g Carotinoid Kristalle erhalten, welche eine
10 Reinheit (HPLC, vgl. oben) von 95 % aufwiesen. Die Ausbeute an Carotinoid Kristallen betrug 80 % bezogen auf die Konzentration an Carotinoid in der Biomasse.

15 Die abgetrennte methylenchloridfeuchte Biomasse wurde nach Wasserdampfdestillation sprühgetrocknet ($T_E = 125\text{ °C}$, $T_A = 60\text{ °C}$) und kann als Tierfuttermitteladditiv eingesetzt werden.

Hierzu wurde nach vorheriger Homogenisation mit einem Dispergiergerät
20 (Ultra-Turrax) und unter ständigem Rühren der Suspension die Zellmasse über eine Schlauchpumpe auf den Trockner aufgegeben.

Die Eindüsung in den Zylinder des Laborsprühtrockners erfolgte dabei über eine Zweistoffdüse mit dem Durchmesser 2,0 mm. Eingedüst wurde
25 mit 2 bar und 4,5 Nm³/h Stickstoff. Die Temperatur am Eintritt betrug ca. 125°C bis 127°C. Das Trocknungsgas war Stickstoff mit einer Flussrate von 22 Nm³/h. Die Austrittstemperatur betrug zwischen 59°C und 61°C. Bei jeder der drei Fermentationsbrühen konnte am Zyklon des Sprüh-
trockners rieselfähiges Produkt abgeschieden werden. Die Wandbeläge im
30 Turm (sofern vorhanden) platzten automatisch von der Gefäßwand ab und wurden als unproblematisch eingestuft.

Es wurden ca. 2,5 – 25 g pulveriges Nahrungsmittel erhalten, welches direkt als Tierfuttermittel verwendet werden könnte. Es enthielt ca. 0,5% - 1,5% Carotinoide bezogen auf das Trockengewicht. Die Restfeuchte betrug weniger als 5%.

Insgesamt (einschließlich des aufgereinigten Carotinoid-Nahrungsmittels) betrug die Ausbeute an Carotinoid ca. 95 % bezogen auf die Ausgangsmenge Carotinoid in der Kulturbrühe.

10

15

20

25

30

Patentansprüche

1. Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderter Organismen der Gattung *Blakeslea* umfassend
- 5
- (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und
- 10
- (vi) Selektion und Vermehrung der gentechnisch veränderten Zelle oder Zellen,
- (vii) Kultivierung der gentechnisch veränderten Zellen,
- (viii) Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe.
- 15
2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es sich um Zellen von Pilzen der Art *Blakeslea trispora* handelt.
3. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector oder freie Nukleinsäuren verwendet werden.
- 20
4. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector in das Genom mindestens einer der Zellen integriert wird.

5. Verfahren nach Anspruch 4, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector einen Promotor und/oder einen Terminator enthält.
6. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 5, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector enthaltend den gpd, pcarB, pcarRA und/oder ptef1 Promotor und/oder den trpC Terminator eingesetzt wird.
7. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 6, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector enthaltend ein Resistenzgen eingesetzt wird.
8. Verfahren nach Anspruch 7, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector ein Hygromycin-Resistenzgen (hph), insbesondere aus E. coli enthält.
9. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der gpd Promotor die Sequenz SEQ ID NO: 1 aufweist.
10. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der trpC Terminator die Sequenz SEQ ID NO: 2 aufweist.
11. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der tef1 Promotor die Sequenz SEQ ID NO: 35 aufweist.
12. Verfahren nach einem der Ansprüche 6 bis 11, **dadurch gekennzeichnet, dass** der gpd Promotor und der trpC Terminator aus Aspergillus nidulans stammen.

13. Verfahren nach einem Ansprüche 3 bis 12, **dadurch gekennzeichnet, dass** der Vector die SEQ ID NO: 3 umfasst.
14. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Transformation (i) mittels Agrobakterien, Konjugation, Chemikalien, Elektroporation, Beschuss mit DNA-beladenen Partikeln, Protoplasten oder Mikroinjektion durchgeführt wird.
15. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in der Homokaryontisierung (ii) ein mutagenes Agens eingesetzt wird.
16. Verfahren nach Anspruch 15, **dadurch gekennzeichnet, dass** als mutagenes Agens N-Methyl-N'-nitro-nitrosoguanidin (MNNG), UV-Strahlung oder Röntgenstrahlung eingesetzt wird.
17. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Selektion durch Markierung und/oder Auswahl der einkernigen Zellen erfolgt.
18. Verfahren nach einem der vorhergehenden Ansprüche 1 - 17, **dadurch gekennzeichnet, dass** in der Selektion 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) oder 5-Fluororotat (FOA) und Uracil und Hygromycin eingesetzt werden.
19. Verfahren nach einem der Ansprüche 3 bis 18, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder deren Vorstufen enthält.
20. Verfahren nach einem der Ansprüche 3 bis 19, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinen oder Xanthophyllen enthält.

21. Verfahren nach einem der Ansprüche 3 bis 20, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, α -Carotin, Lutein, Phytofluen, Bixin oder Phytoen enthält.

10

22. Verfahren zur Bereitstellung mindestens eines hochreinen Carotinoids und eines Nahrungsmittels, enthaltend Carotinoide-produzierende Organismen und mindestens das eine Carotinoid, umfassend nach der Kultivierung von Carotinoide-produzierenden Organismen der Gattung *Blakeslea* die Schritte

15

- I) Abtrennung der Biomasse,
 - IA) ggf. Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser,
 - IB) Sterilisation und Zellaufschluß der Biomasse,
 - IC) ggf. Trocknung und/oder homogene Verteilung und
- II) partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse,
 - IIA)

20

25

- 1) Entfernung von Lösemittelresten aus der Carotinoid-haltigen Biomasse,
- 2) ggf. homogene Suspension der Biomasse mit einem Biomasse-Feststoffgehalt > 10
- 5 3) Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels,

IIB)

- 1) Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.
- 10

23. Verfahren nach Anspruch 22, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid aus der Gruppe bestehend aus Carotinen und Xanthophyllen ausgewählt ist.

15 24. Verfahren nach Anspruch 22 oder 23, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid aus der Gruppe bestehend aus Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, Lutein, Phytofluen, Bixin und Phytoen ausgewählt ist.

20 25. Verfahren nach einem der Ansprüche 22 bis 24, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid Astaxanthin, Zeaxanthin, *Bixin* oder Phytoen ist.

26. Verfahren nach einem der Ansprüche 22-25, **dadurch gekennzeichnet, dass** die Sterilisation und der Zellaufschluß mittels Wasserdampf oder Mikrowellenstrahlung durchgeführt werden.

25

27. Verfahren nach einem der Ansprüche 22-26, **dadurch gekennzeichnet, dass** die Extraktion der Carotinoide aus der Biomasse mittels Me-

thylenchlorid oder überkritischem Kohlendioxid oder Tetrahydrofuran durchgeführt wird.

28. Verfahren nach Anspruch 27, **dadurch gekennzeichnet, dass** die im überkritischen Kohlendioxid gelösten Carotinoide direkt isoliert werden
5 oder in Methylenchlorid aufgenommen werden.
29. Verfahren nach einem der Ansprüche 22-28, **dadurch gekennzeichnet, dass** die Extraktion der Carotinoide aus der Biomasse ein oder ggf. mehrstufig erfolgt.
30. Verfahren nach einem der Ansprüche 22-29, **dadurch gekennzeichnet, dass** die Entfernung von Lösungsmitteln aus der Biomasse im
10 Schritt IA1) mittels Wasserdampf-Destillation .
31. Verfahren nach einem der Ansprüche 22-30, **dadurch gekennzeichnet, dass** die Trocknung in Schritt IIA3) mittels Sprühtrocknung oder Kontakttrocknung durchgeführt wird.
- 15 32. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Kristallisation im Schritt IIB1) durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel erfolgt.
33. Verfahren nach Anspruch 32, **dadurch gekennzeichnet, dass** der
20 Austausch des verwendeten Lösungsmittels gegen Wasser oder einen niederen Alkohol, insbesondere Methanol erfolgt.
34. Verfahren nach Anspruch 13, **dadurch gekennzeichnet, dass** der gentechnisch veränderte Organismus der Gattung *Blakeslea* durch Transformation mit einem Vector, der eine Sequenz aus der Gruppe
25 bestehend aus den SEQ ID NO: 37 – 51 und 62 aufweist, herstellbar ist.

35. Verfahren zur Herstellung eines Nahrungsmittels enthaltend Organismen der Gattung *Blakeslea* und mindestens ein Carotinoid, umfassend nach der Kultivierung von Carotinoide-produzierenden Organismen der Gattung *Blakeslea* die Schritte

- 5 I) Homogene Suspendierung der Feststoffe der Kulturbrühe und
- IIA) bei einem Biomasse-Feststoffgehalt der Kulturbrühe von > 2 %
- 10 1) ggf. Konzentration der Kulturbrühe auf einen Feststoffgehalt < 50 % und
- 2) Trocknung der Kulturbrühe zur Herstellung des Nahrungsmittels
- oder
- 15 IIB) bei einem Feststoffgehalt von < 2 % der Kulturbrühe,
- 1) Konzentration der Kulturbrühe auf einen Feststoffgehalt > 2 % und < 50 % und
- 2) Trocknung der Suspension zur Herstellung des Nahrungsmittels,
- 20 oder
- IIC) unabhängig vom Feststoffgehalt der Kulturbrühe,
- 1) Abtrennung der Biomasse,
- 25 2) ggf. Waschen der Biomasse mit Carotinoide nicht lösenden Lösungsmitteln, insbesondere Wasser,
- 3) Sterilisation und Zellaufschluß,
- 4) ggf. Trocknung und homogene Verteilung,
- 5) partielle Extraktion der Carotinoide aus der Biomasse
- 30 mittels eines Carotinoide lösendes Lösungsmittels,

- 5a) Abtrennung der Carotinoid-haltigen Biomasse vom Carotinoid-haltigen Lösungsmittel,
- 5b) Entfernung von Lösemittelresten aus der Biomasse und
- 5c) Trocknung der Biomasse zur Herstellung des Nahrungsmittels,
- 6) Kristallisation der Carotinoide aus dem in 5a) verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.
- 10 36. Verfahren nach Anspruch 35, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid aus der Gruppe bestehend aus Carotinen und Xanthophyllen ausgewählt ist.
- 15 37. Verfahren nach Anspruch 35 oder 36, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid aus der Gruppe bestehend aus Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Anthonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, Lutein, Bixin, Phytoen ausgewählt ist.
- 20 38. Verfahren nach einem der Ansprüche 35-37, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid Astaxanthin, Zeaxanthin, *Bixin* oder Phytoen ist.
39. Verfahren nach einem der Ansprüche 35-38, **dadurch gekennzeichnet, dass** die Sterilisation und der Zellaufschluß im Schritt II3) mittels Wasserdampf oder Mikrowellenstrahlung durchgeführt wird.
- 25 40. Verfahren nach einem der Ansprüche 35-39, **dadurch gekennzeichnet, dass** die Extraktion der Carotinoide aus der Biomasse im Schritt IIC5) mittels Methylenchlorid oder überkritischen Kohlendioxid durchgeführt wird.

41. Verfahren nach Anspruch 40, **dadurch gekennzeichnet**, dass die im überkritischen Kohlendioxid gelösten Carotinoide direkt isoliert werden oder in Methylenchlorid aufgenommen werden..
- 5 42. Verfahren nach einem der Ansprüche 35-41, **dadurch gekennzeichnet**, dass die Extraktion der Carotinoide aus der Biomasse ein- oder ggf. mehrstufig erfolgt.
43. Verfahren nach einem der Ansprüche 35-42, **dadurch gekennzeichnet**, dass die Entfernung von Lösungsmitteln aus der Biomasse im Schritt IIC5b) mittels Wasserdampf-Destillation.
- 10 44. Verfahren nach einem der Ansprüche 35-43, **dadurch gekennzeichnet**, dass die Trocknung in einem der Schritte IIA1), IIB2) oder IIC5c) mittels Sprühtrocknung oder Kontakt durchgeführt wird.
- 15 45. Verfahren nach einem der Ansprüche 35-44, **dadurch gekennzeichnet**, dass die Kristallisation im Schritt IIC6) durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel erfolgt.
46. Verfahren nach Anspruch 45, **dadurch gekennzeichnet**, dass der Austausch des verwendeten Lösungsmittels gegen Wasser oder einen niederen Alkohol, insbesondere Methanol erfolgt.
- 20 47. Verfahren nach einem der Ansprüche 35-46, **dadurch gekennzeichnet**, dass der gentechnisch veränderte Organismus der Gattung *Blakeslea* durch Transformation mit einem Vector, der eine Sequenz aus der Gruppe bestehend aus den SEQ ID NO: 37 – 51 und 62 aufweist, herstellbar ist.
- 25 48. Nahrungsmittel, insbesondere Tierfuttermittel herstellbar nach einem der Verfahren der Ansprüche 1 bis 47.

49. Nahrungsergänzungsmittel, insbesondere Tierfutterergänzungsmittel herstellbar nach einem der Verfahren der Ansprüche 1 bis 47.

50. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß Nahrungsmittel und Tierfuttermittel aus einer Fermentation erhältlich sind.

51. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß Nahrungsergänzungsmittel und Tierfutterergänzungsmittel aus einer Fermentation erhältlich sind.

52. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß mindestens zwei Produkte aus der Gruppe Nahrungsmittel, Nahrungsergänzungsmittel, Tierfuttermittel und Tierfutterergänzungsmittel aus einer Fermentation erhältlich sind.

53. Verwendung der nach einem der Verfahren der Ansprüche 1 bis 14 erhältlichen Carotinoide zur Herstellung von kosmetischen, pharmazeutischen, dermatologischen Zubereitungen, Nahrungsmitteln, Nahrungsergänzungsmitteln, Tierfuttermitteln oder Tierfutterergänzungsmitteln.

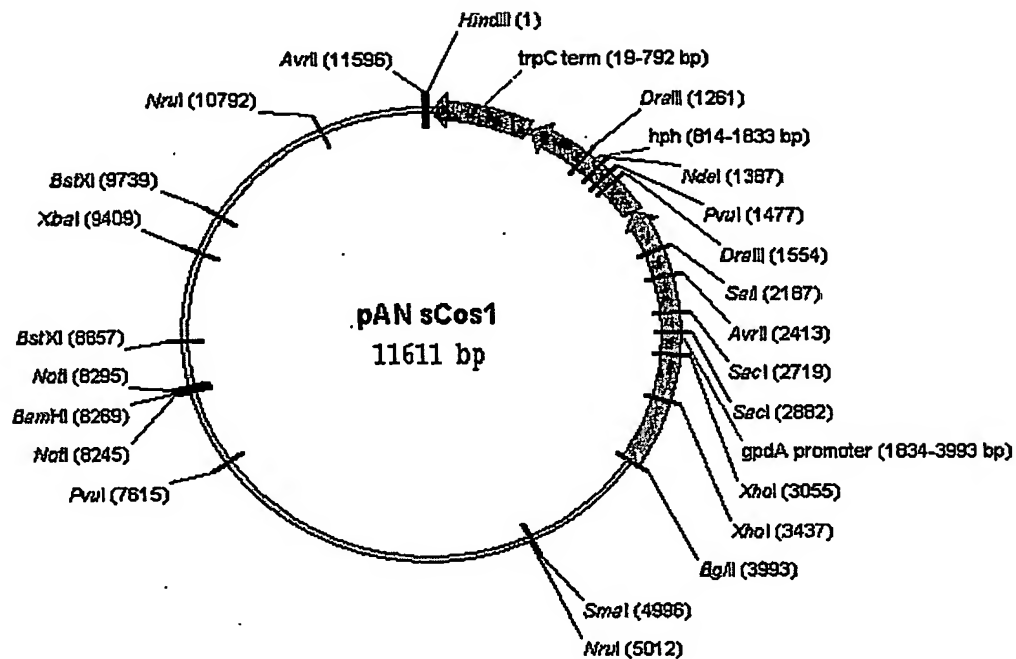
Fig. 1: Vektor pANsCos1

Fig. 2: Vektor pBinAHyg

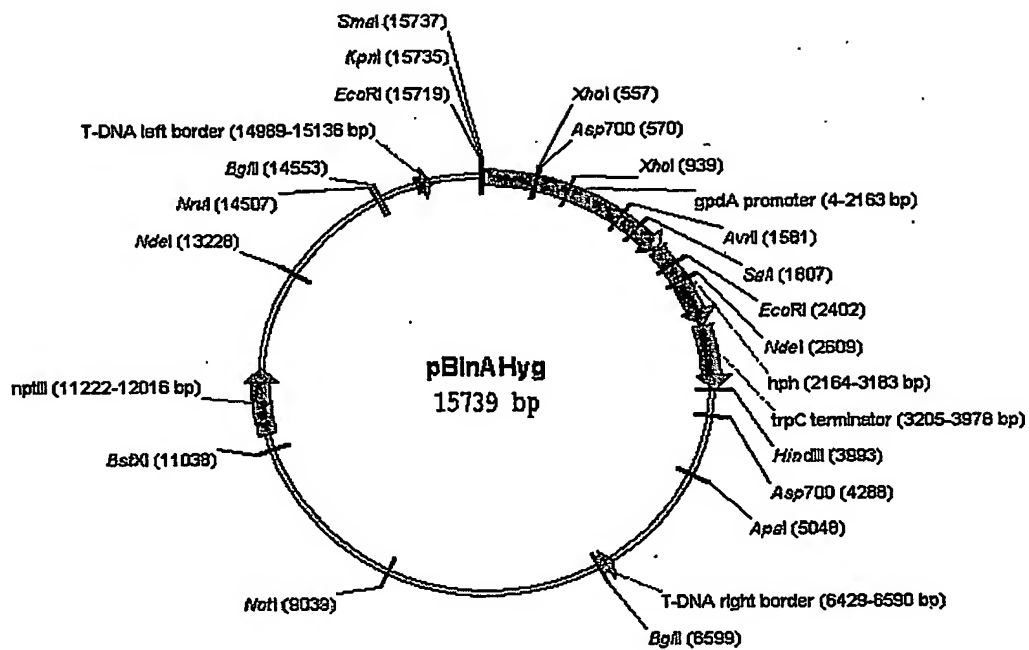


Fig. 3: Gels des Ergebnis einer PCR
Spur:

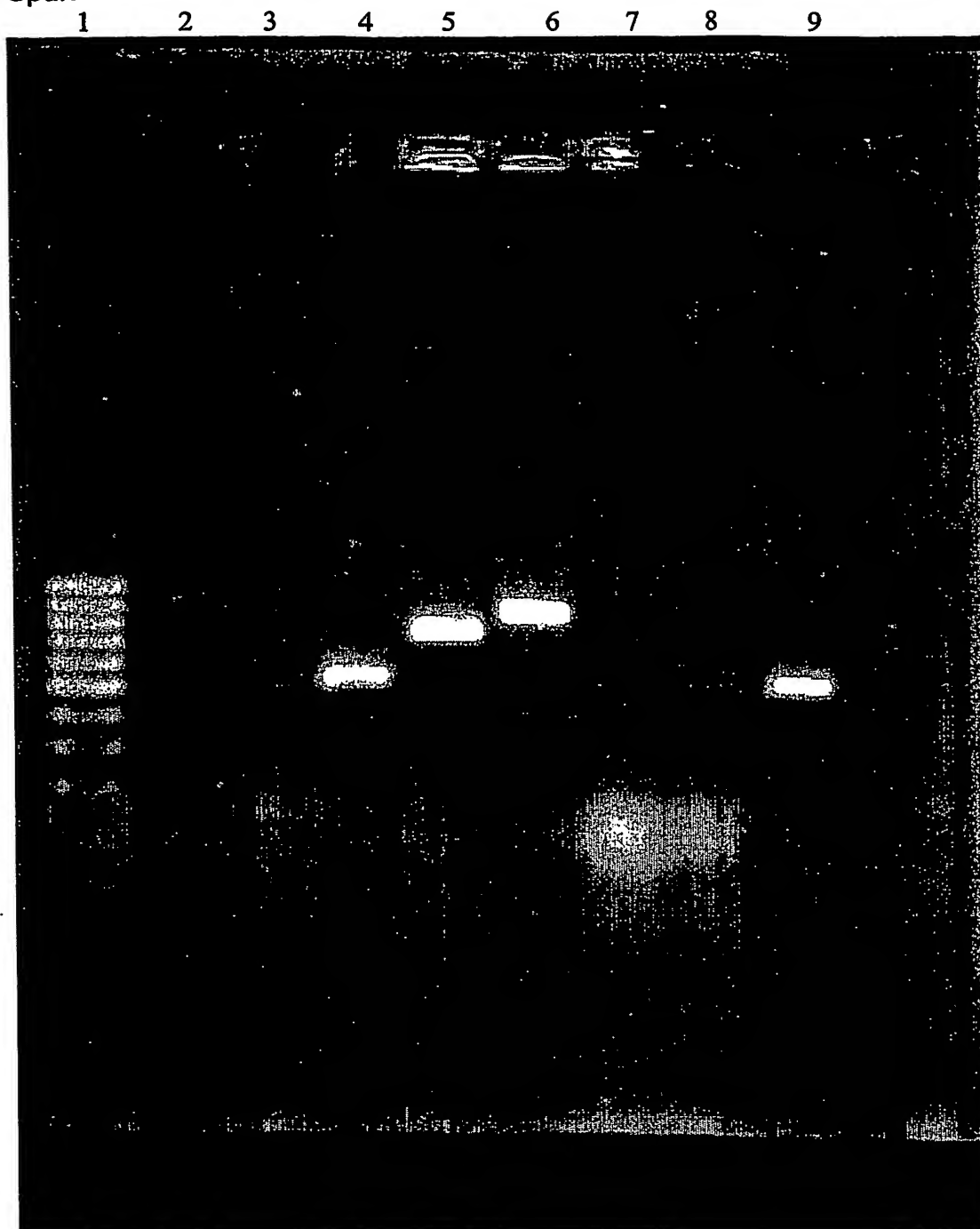


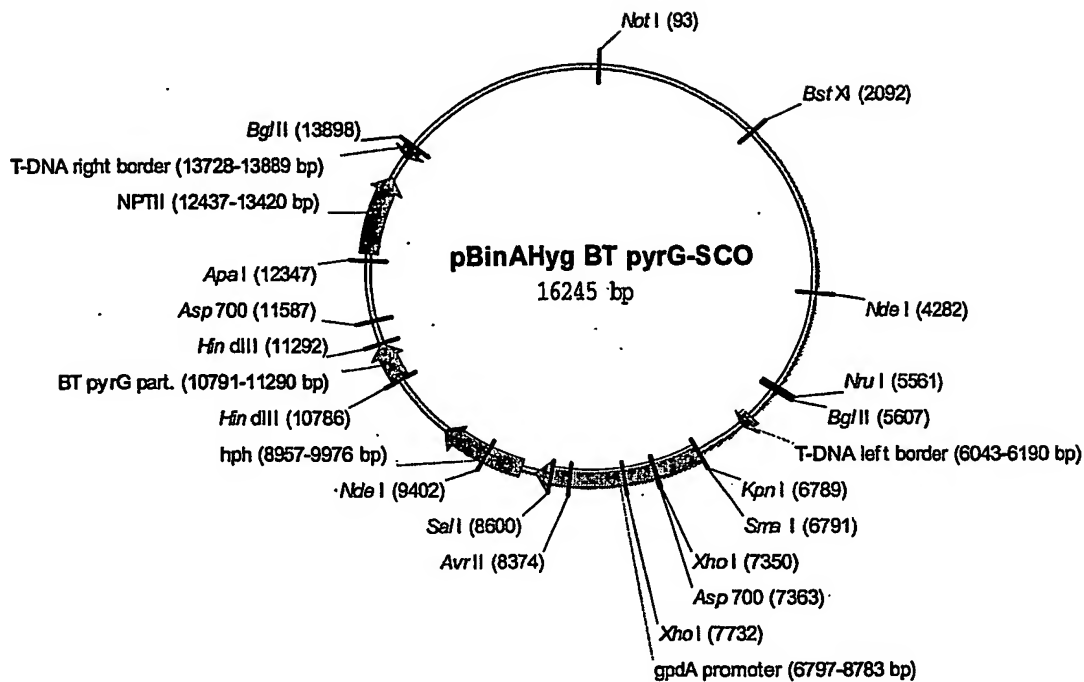
Fig. 4: Plasmid pBinAHygBTpyrG-SCO

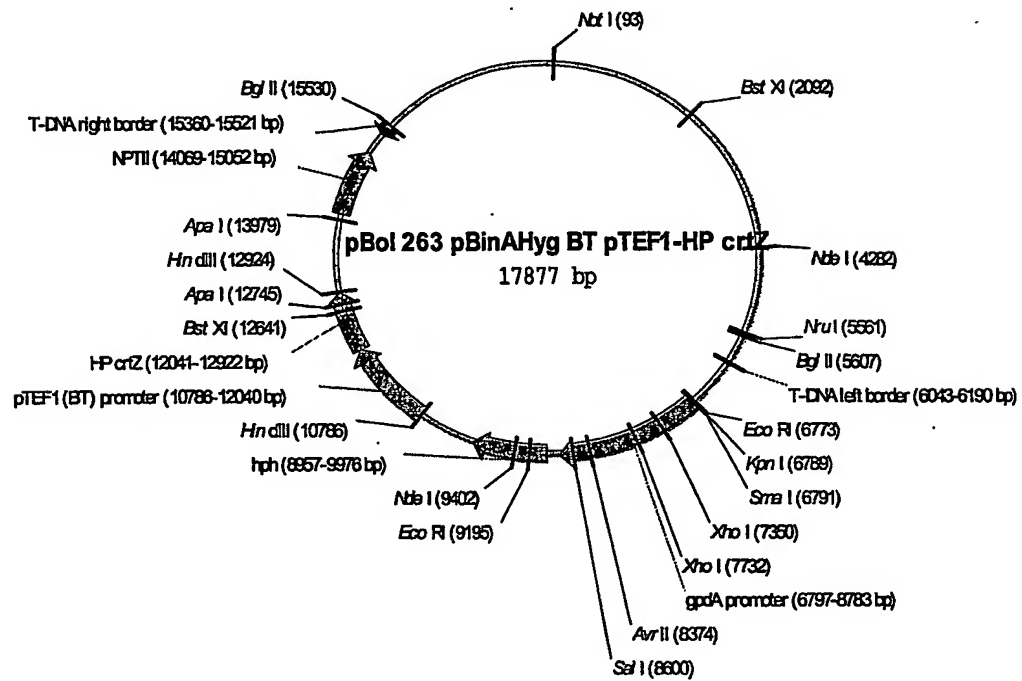
Fig. 5: Plasmid pBinAHygBTpTEF1-HPcrtZ

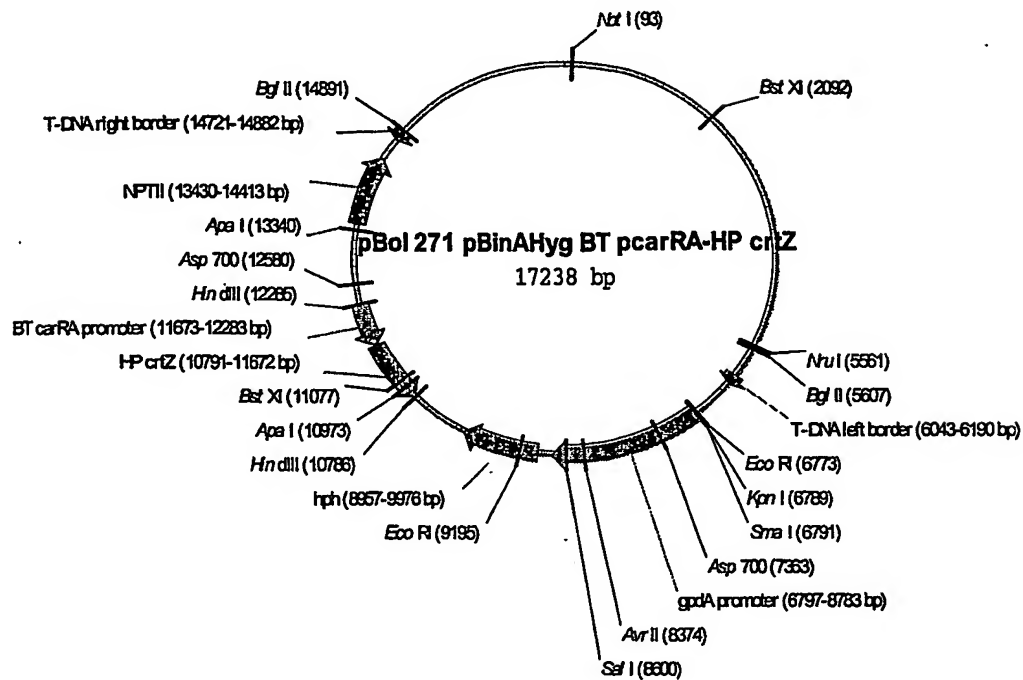
Fig. 6: Plasmid pBinAHyg-BTpcarRA-HPcrtZ

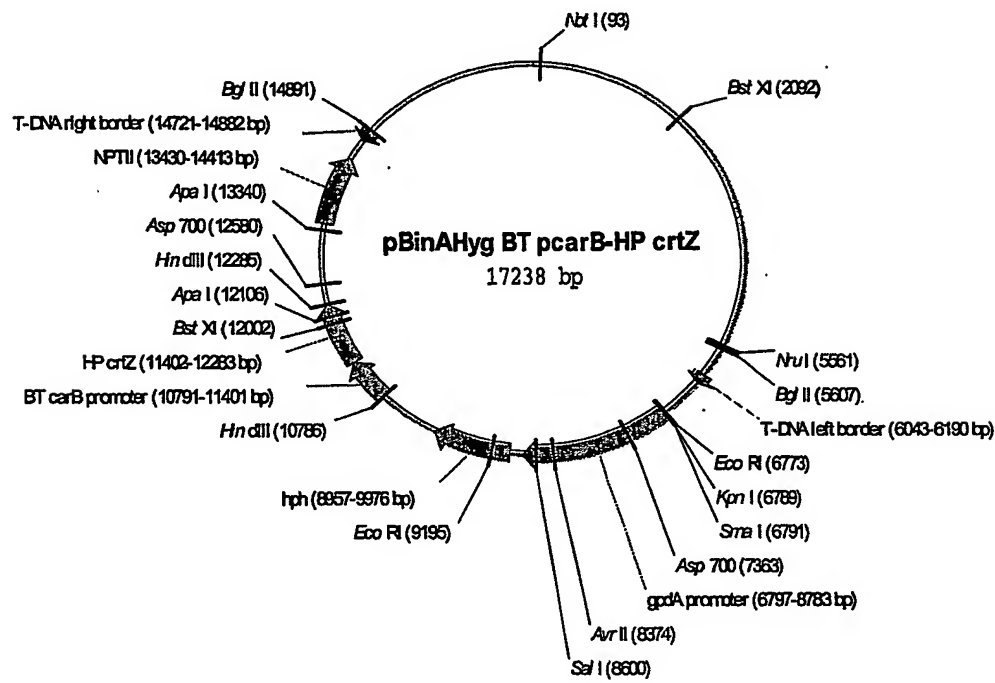
Fig. 7: Plasmid pBinAHygBTpcarB-HPcrtZ

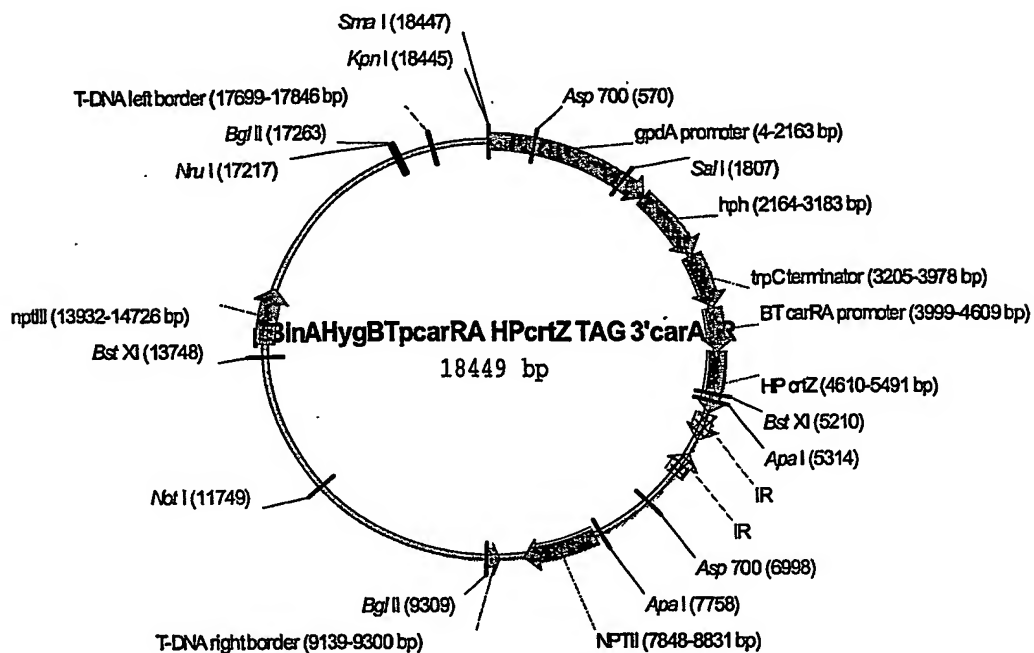
Fig. 8: Plasmid p-carRA-HPcrtZ-TAG-3'carA-IR

Fig. 9: Plasmid p-carRA-HPcrtZ-GCG-3'carA-IR

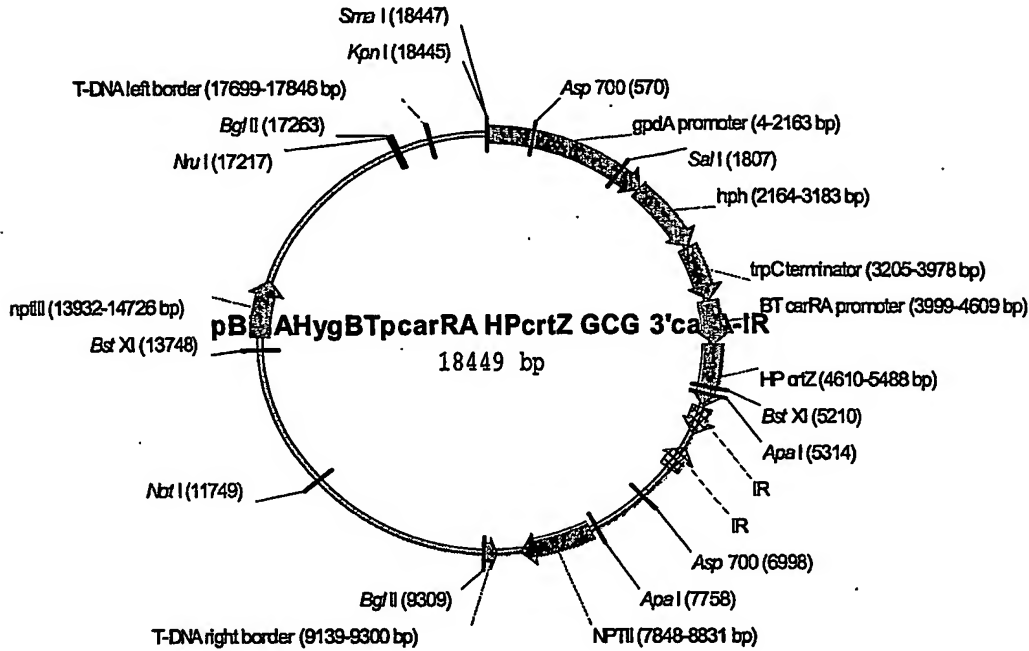


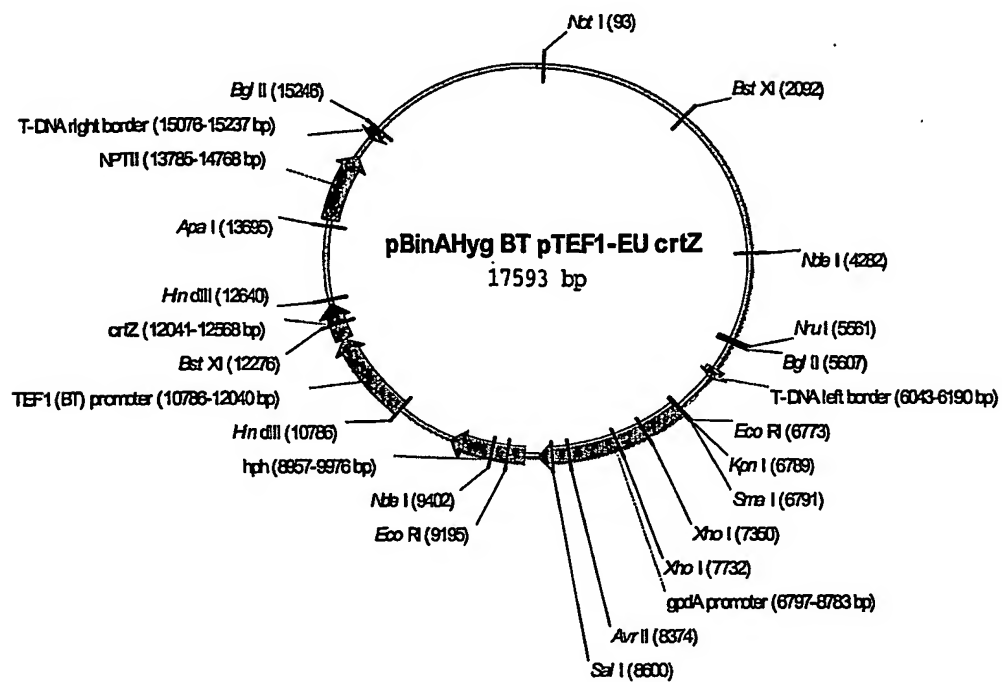
Fig. 10: Plasmid pBinAHygBTpTEF1-EUcrtZ

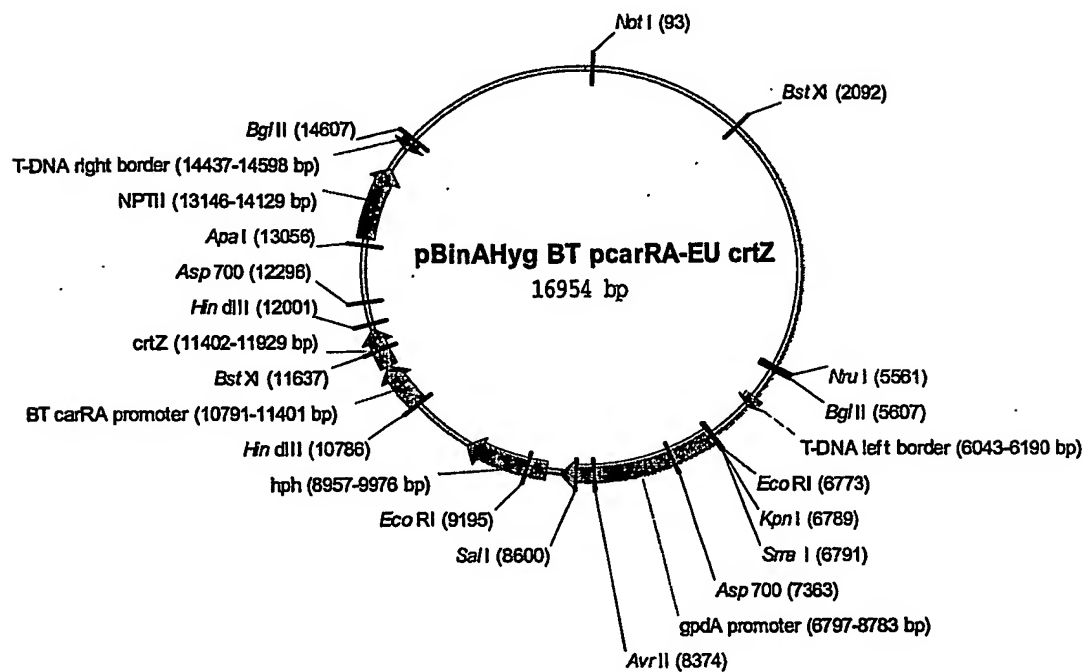
Fig. 11: Plasmid pBinAHygBTpcarRA-EUcrtZ

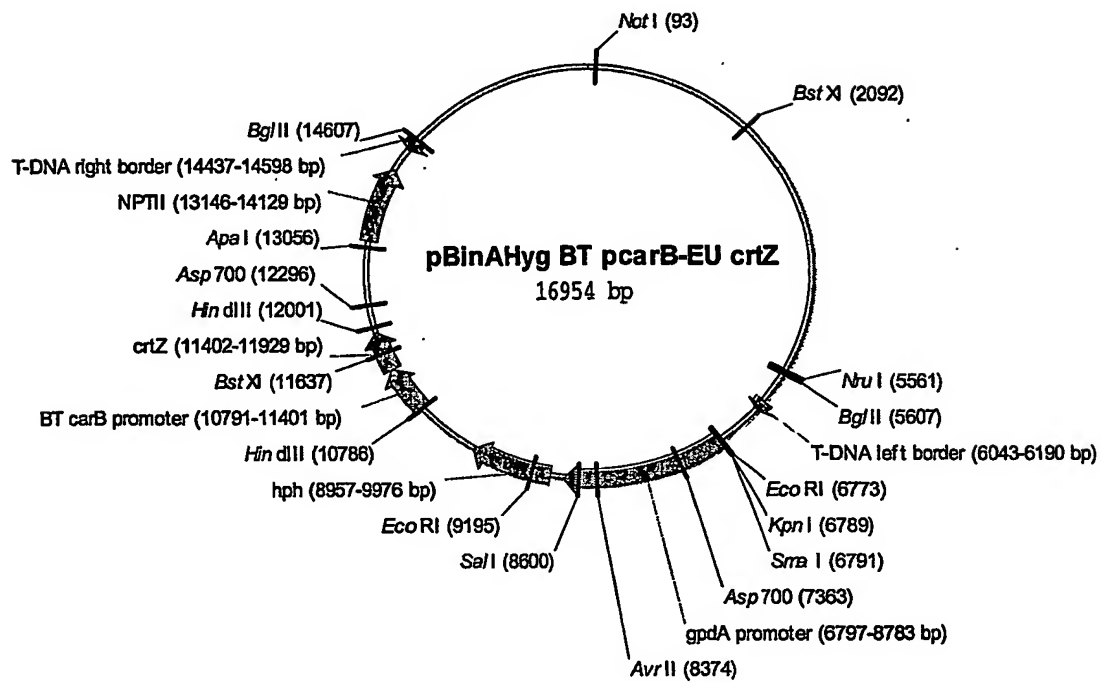
Fig. 12: Plasmid pBinAHygBTpcarB-EUcrtZ

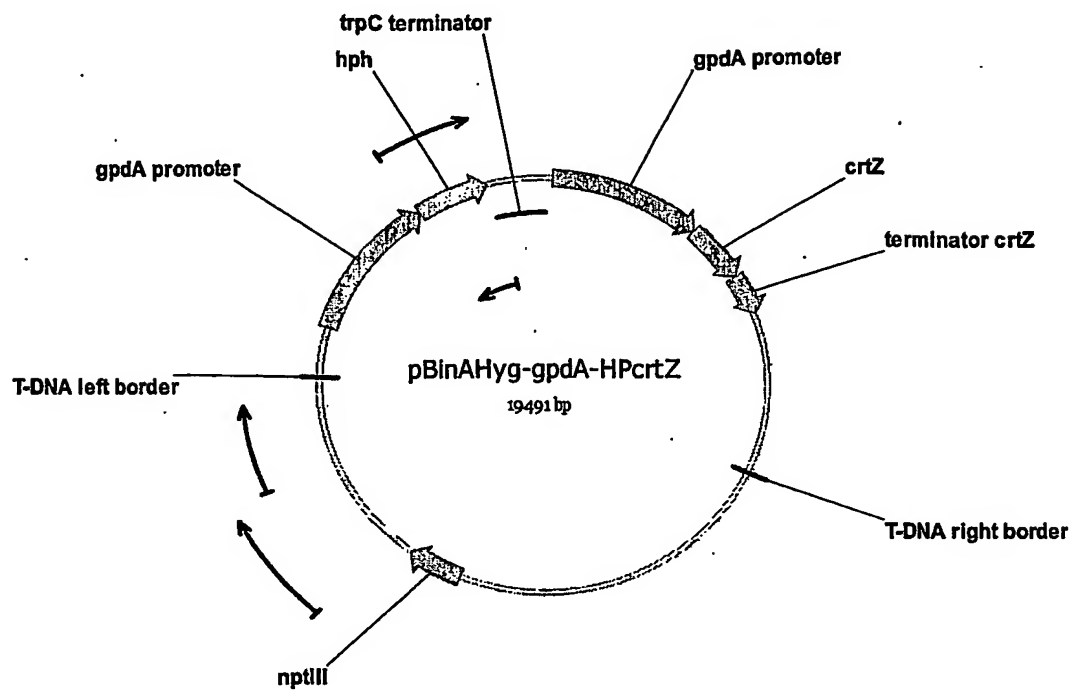
Fig. 13: Plasmid p-BinAHyg-gpdA-HPcrtZ

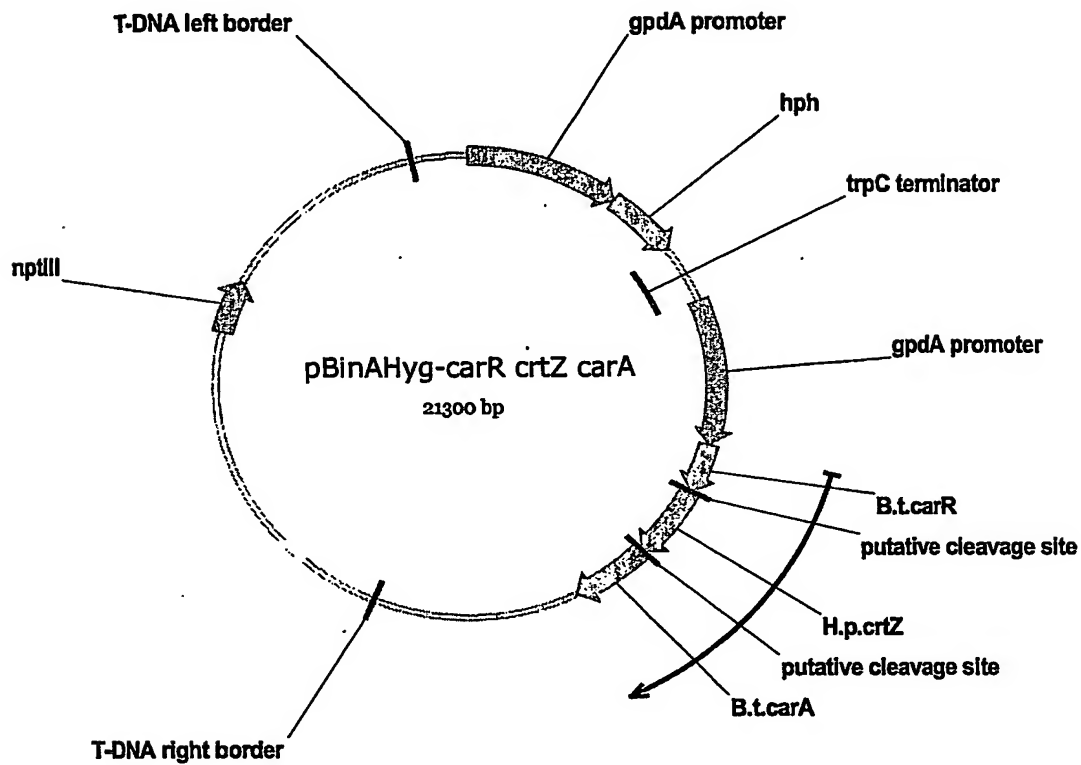
Fig. 14: Plasmid pBinAHyg-carRcrtZcarA

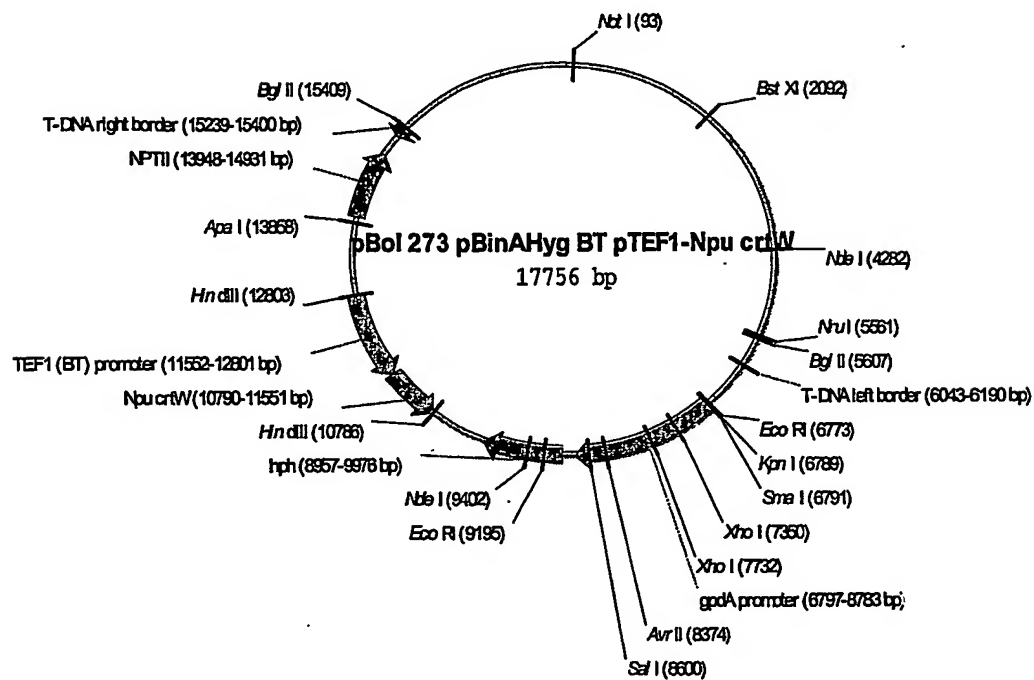
Fig. 15: Plasmid pBinAHyg-BTpTEF1-NPcrtW

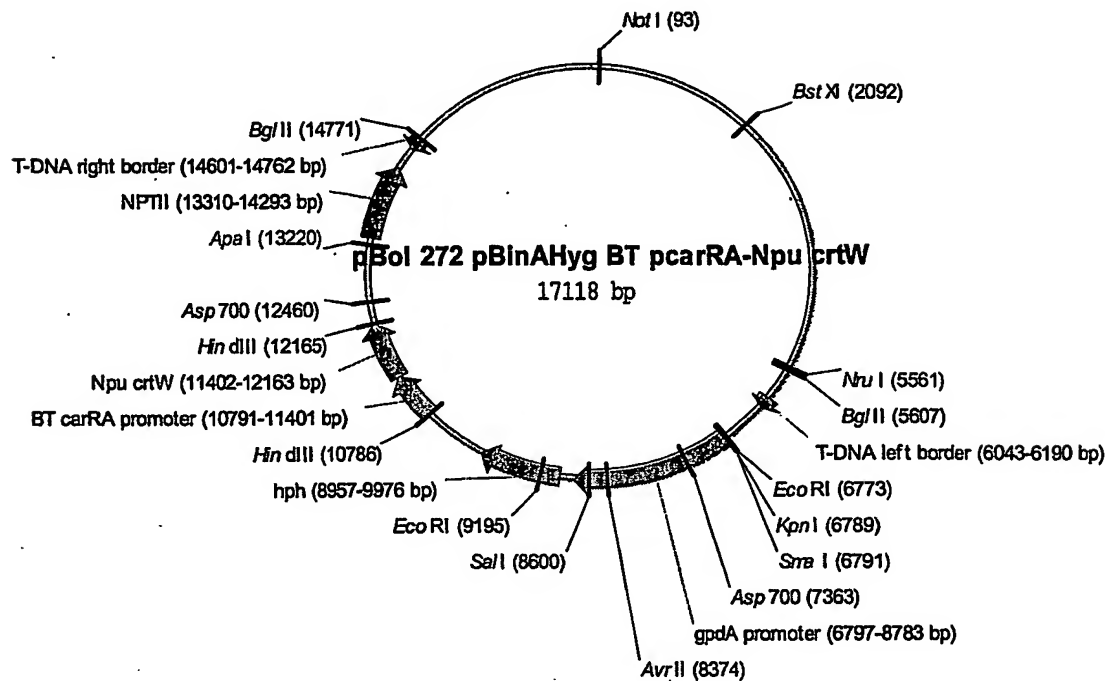
Fig. 16: Plasmid pBinAHyg_BTpcarRA_NPcrtW

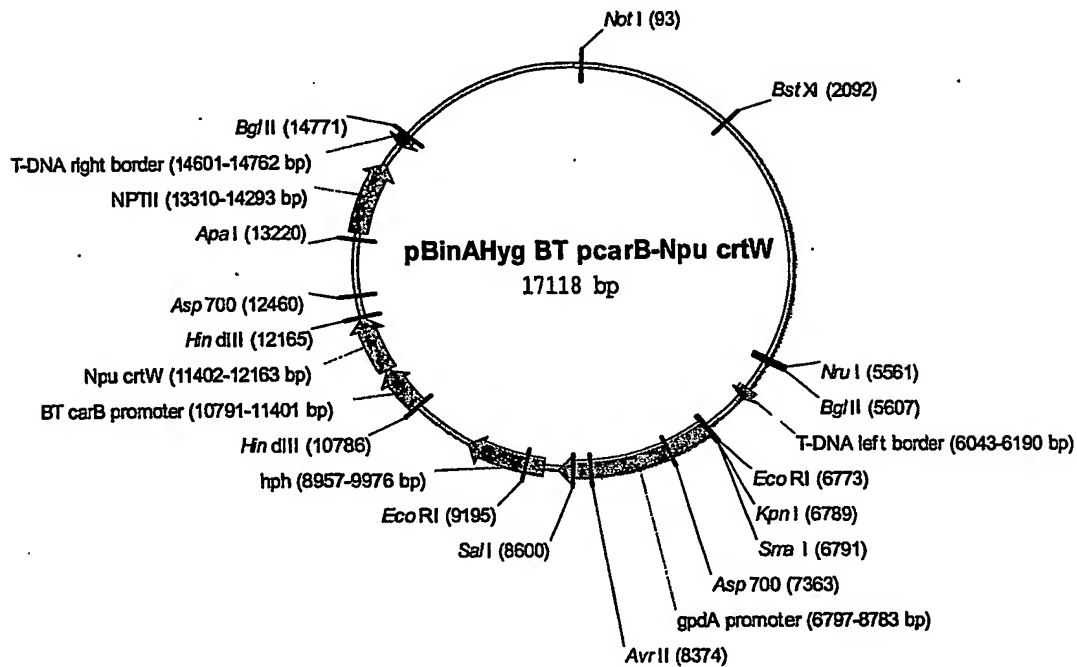
Fig. 17: Plasmid pBinAHyg-BTpcarB-NPcrtW

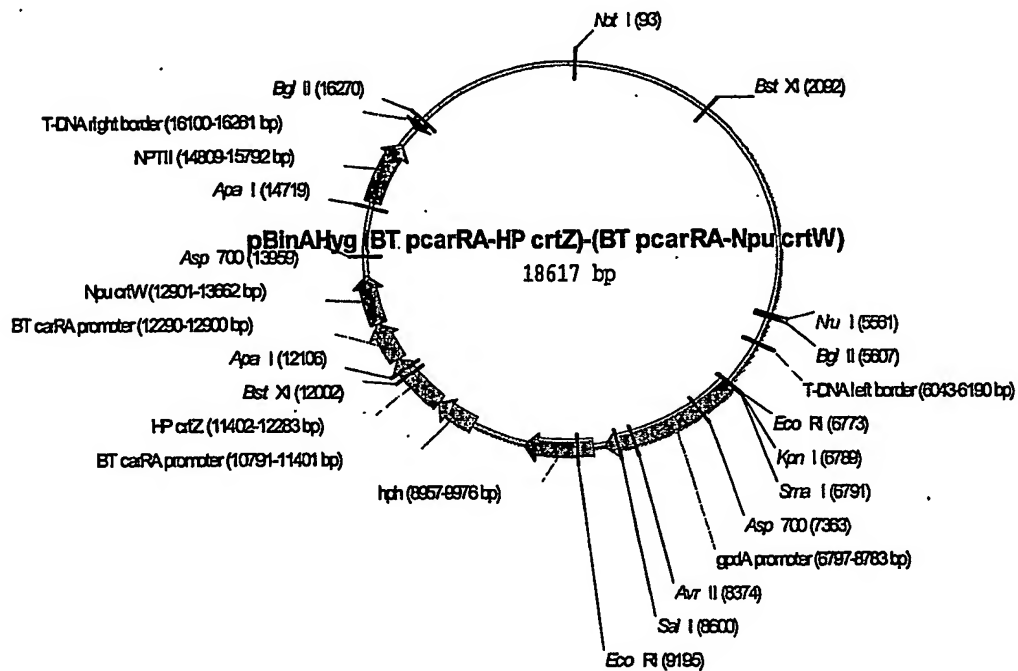
Fig. 18: Plasmid pBinAHygBTpcarRA-HPcrtZ-BTpcarRA-Npu crtW

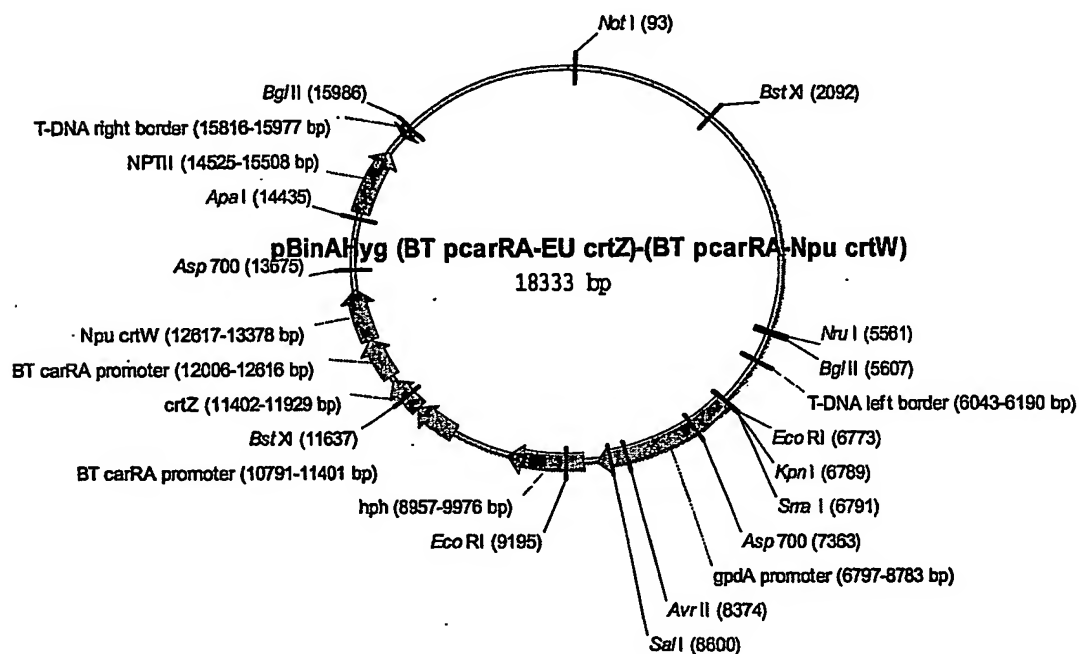
Fig. 19: Plasmid pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW

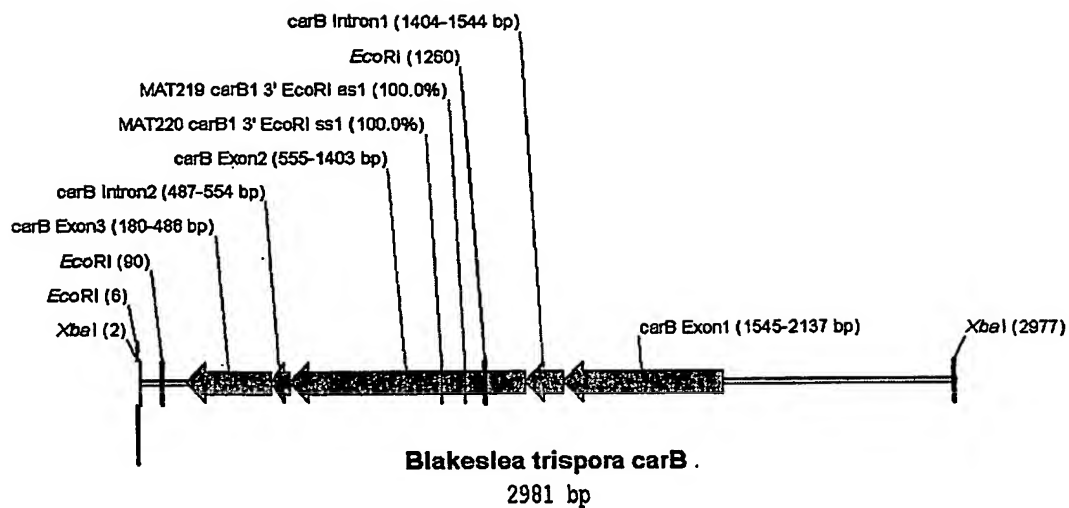
Fig. 20: carB

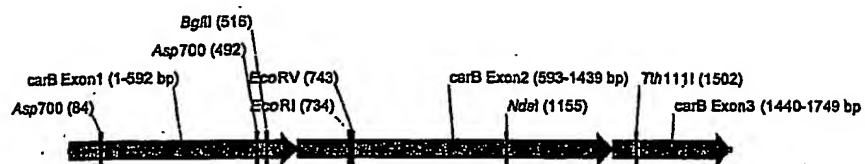
Fig. 21: CDS von carB

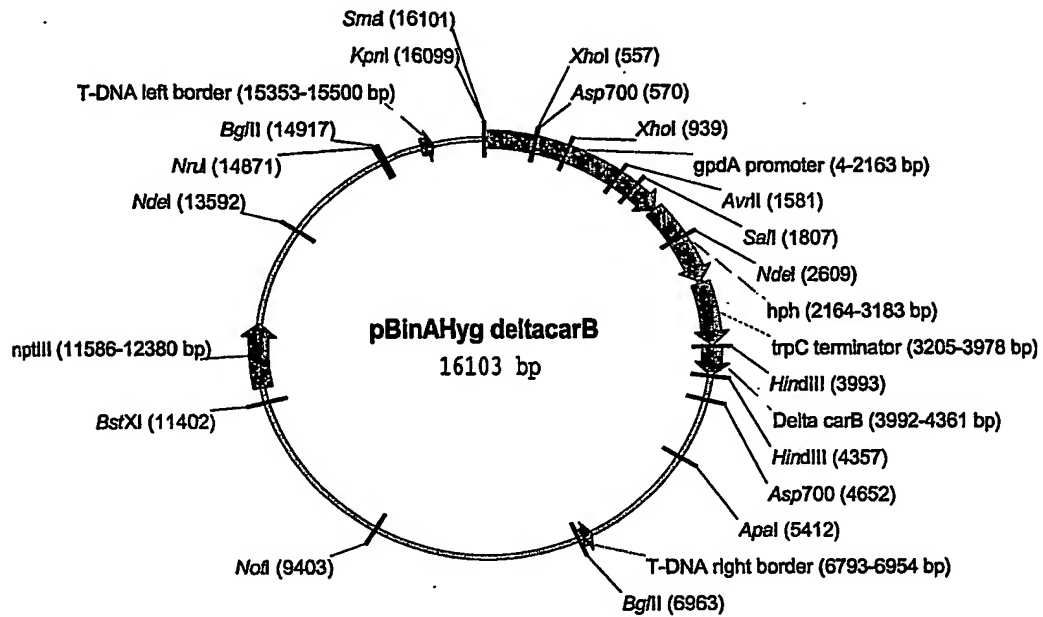
Fig. 22: Vektor pBinAHyg Δ carB

Fig. 23: HPLC Standard

Sequence File : C:\HPCHEM\2\SEQUENCE\CAROTIN.S
Method : C:\HPCHEM\2\METHODS\CAROTIN.M
Last changed : 11/12/2003
Carotinocide Nucleosil 100-7 μ C18 250*4, 0mm
DAD1 A, Slg=450,20 Ret=off (131103.1VPROBE018.D)

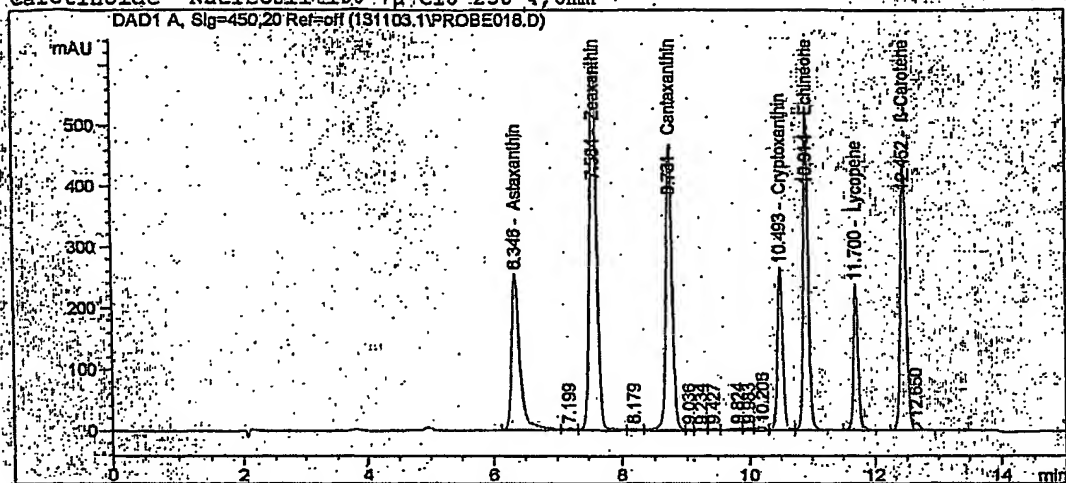
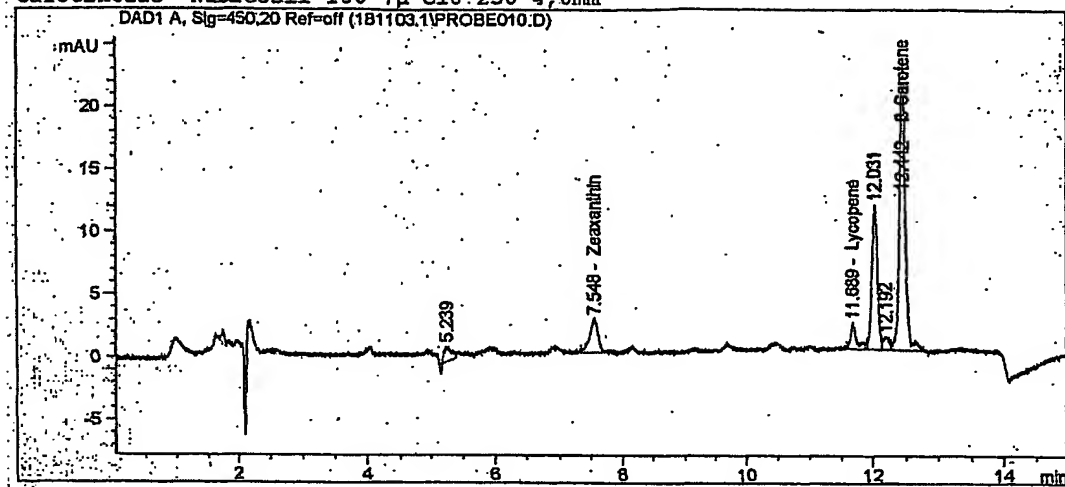


Fig. 24: HPLC

Sequence File : C:\HPCHEM\2\SEQUENCE\CAROTIN.S
Method : C:\HPCHEM\2\METHODS\CAROTIN.M
Last changed : 11/18/2003
Carotinoide Nucleosil 100-7µ C18.250*4,0mm



SEQUENCE LISTING

<110> BASF AG

<120> Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen
mittels gentechnisch veränderter Organismen der Gattung
Blakeslea, mit dem Verfahren hergestellte Carotinoide oder deren
Vorstufen und deren Verwendung

<130> BASF/NAE877/03

<160> 80

<170> PatentIn version 3.2

<210> 1

<211> 2160

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 1

ctttcgacac tgaaatacgt cgagcctgct ccgcttgaa gcggcgagga gcctcgtcct	60
gtcacaacta ccaacatgga gtacgataag ggccagttcc gccagctcat taagagccag	120
ttcatgggcg ttggcatgat ggccgcatg catctgtact tcaagtacac caacgctctt	180
ctgatccagt cgatcatccg ctgaaggcgc tttcgaatct ggtaagatc cacgtcttcg	240
ggaagccagc gactggtgac ctccagcgtc cctttaaggc tgccaacagc tttctcagcc	300
agggccagcc caagaccgac aaggcctccc tccagaacgc cgagaagaac tggaggggtg	360
gtgtcaagga ggagtaagct ccttattgaa gtcggaggac ggagcgggtg caagaggata	420
ttcttcgact ctgtattata gataagatga tgaggaattg gaggtagcat agcttcattt	480
ggatttgctt tccaggctga gactctagct tggagcatag agggtccttt ggctttcaat	540

attctcaagt atctcgagtt tgaacttatt ccctgtgaac cttttattca ccaatgagca . 600

ttggaatgaa catgaatctg aggactgcaa tcgccatgag gttttcgaaa tacatccgga . 660

tgtcgaaggc ttggggcacc tgcgttggtt gaatttagaa cgtggcacta ttgatcatcc 720

gatagctctg caaagggcgt tgcacaatgc aagtcaaacg ttgctagcag ttccagggtgg 780

aatgttatga tgagcattgt attaaatcag gagatatagc atgatctcta gttagctcac 840

cacaaaagtc agacggcgta accaaaagtc acacaacaca agctgtaagg atttcggcac 900

ggctacggaa gacggagaag ccaccttcag tggactcgag taccatttaa ttctatttgt 960

gtttgatcga gacctaatac agcccctaca acgaccatca aagtcgtata gctaccagtg 1020

aggaagtgga ctcaaatacga cttcagcaac atctcctgga taaactttaa gcctaaacta 1080

tacagaataa gatagggtgga gagcttatac cgagctccca aatctgtcca gatcatgggt 1140

gaccggtgcc tygatcttcc tatagaatca tccttattcg ttgacctagc tgattctgga 1200

gtgaccaga gggatcatgac ttgagcctaa aatccgccgc ctccaccatt tgtagaaaaa 1260

tgtgacgaac tcgtgagctc tgtacagtga ccggtgactc tttctggcat gcggagagac 1320

ggacggacgc agagagaagg gctgagtaat aagccactgg ccagacagct ctggcggctc 1380

tgagggtcag tggatgatta ttaatccggg accggccgcc cctccgcccc gaagtggaaa 1440

ggctggtgtg cccctcgttg accaagaatc tattgcatca tcggagaata tggagcttca 1500

tcgaatcacc ggcagtaagc gaaggagaat gtgaagccag ggggtgtatag ccgtcggcga 1560

aatagcatgc cattaaccta ggtacagaag tccaattgct tccgatctgg taaaagattc 1620

acgagatagt accttctccg aagtaggtag agcgagtacc cggcgcgtaa gctccctaata 1680

tggcccatcc ggcattctgta gggcgtccaa atatcgtgcc tctcctgctt tgcccgggtg 1740

atgaaaccgg aaaggccgct caggagctgg ccagcggcgc agaccgggaa cacaagctgg 1800

cagtcgaccc atccggtgct ctgcactcga cctgctgagg tccctcagtc cctggtaggc 1860
agctttgccc cgtctgtccg cccggtgtgt cggcgggggtt gacaaggtcg ttgcgtcagt 1920
ccaacatttg ttgcatatt ttctgtctct cccaccagc tgctcttttc tttctctttt 1980
cttttcccat cttcagtata ttcatcttcc catccaagaa cttttatttc ccctaagtaa 2040
gtactttgct acatccatac tccatccttc ccatccctta ttcttttgaa cttttcagtt 2100
cgagcttttc cacttcatcg cagcttgact aacagctacc ccgcttgagc agacatcacc 2160

<210> 2

<211> 774

<212> DNA

<213> Artificial

<220>

<223> Terminator

<220>

<221> misc_feature

<222> (267)..(267)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (475)..(475)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (566)..(566)

<223> n is a, c, g, or t

<400> 2

cgatccactt aacgttactg aaatcatcaa acagcttgac gaatctggat ataagatcgt 60
tggtgtcgat gtcagctccg gagttgagac aaatgggtgtt caggatctcg ataagatacg 120

ttcatttgtc caagcagcaa agagtgcctt ctagtgattt aatagctcca tgtcaacaag 180
aataaaacgc gttttcgggt ttacctcttc cagatacagc tcatctgcaa tgcattaatg 240
cattgactgc aacctagtaa cgccttncag gctccggcga agagaagaat agcttagcag 300
agctattttc attttcggga gacgagatca agcagatcaa cggtcgtcaa gagacctacg 360
agactgagga atccgctctt ggctccacgc gactatatat ttgtctctaa ttgtactttg 420
acatgctcct cttctttact ctgatagctt gactatgaaa attccgtcac cagcncctgg 480
gttcgcaaag ataattgcat gtttcttcct tgaactctca agcctacagg acacacattc 540
atcgtaggta taaacctcga aatcanttcc tactaagatg gtatacaata gtaaccatgc 600
atgggttgct agtgaatgct ccgtaacacc caatacgccg gccgaaactt tttacaact 660
ctcctatgag tcgtttaccc agaatgcaca ggtacacttg tttagaggta atccttcttt 720
ctagctagaa gtcctcgtgt actgtgtaag cgcccactcc acatctccac tcga 774

<210> 3

<211> 15739

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 3

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt	60
cctgtcacia ctaccaacat ggagtagcat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctgggtaag atccacgtct	240
tccggaagcc agcgactggg gacctccagc gtccctttaa ggctgccaac agctttctca	300
gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggg cgttgcacia tgcaagtcaa acgttgctag cagttccagg	780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg	900
cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt	960
tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca	1020
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa	1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcatccttat tcgttgacct agctgattct 1200

ggagtgaccc agaggggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gcccctccgc cccgaagtgg 1440

aaaggctggt gtgccccctg ttgaccaaga atctattgca tcatcggaga atatggagct 1500

tcatcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgac tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccgg 1740

tgtatgaaac cggaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctgt ccgcccgggtg tgcggcggg gttgacaagg tcgttgcgtc 1920

agtccaacat ttgttgccat attttctctg tctccccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccaactca tcgcagcttg actaacagct acccggcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctcgaggggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcggtg gatatgtcct gcgggtaaata agctgcgccg atgggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgc tgacattggg 2400

gaattcagcg agagcctgac ctattgcac tcccgccgtg cacagggtgt cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520

atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580

ggccaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640

tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg cccgaagtc cggcacctcg tgcacgcgga ttccggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcggta ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgacgg atcgccgcgg 2940

ctccgggcgt atatgctccg cattgggtctt gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtccg atccggagcc 3060

gggactgtcg ggcgtacaca aatcgcccgc agaagcgcgg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cggcccagca ctcgctccgag ggcaaaggaa 3180

tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ctttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat ttccattttc gggagacgag atcaagcaga 3540

tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600

atattttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780

gatggtatac aatagtaacc atgcatgggt gcctagttaa tgctccgtaa cacccaatac 3840

gccggccgaa actttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900

cttgttttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct ccactcgacc tgcaggcatg caagcttggc gtaatcatgg tcatagctgt 4020

ttctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcataa 4080

agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac 4140

tgccccgttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 4200

cggggagagg cggtttgcgt attggggcaa agacaaaagg gcgacattca accgattgag 4260

ggagggaagg taaatattga cggaattat tcattaaagg tgaattatca ccgtcaccga 4320

cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4380

ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat 4440

caagtttgcc tttagcgta gactgtagcg cgttttcatc ggcattttcg gtcatagccc 4500

ccttattagc gtttgccatc ttttcaaat caaaatcacc ggaaccagag ccaccaccgg 4560

aaccgcctcc ctgagagccg ccaccctcag aaccgccacc ctgagagcca ccaccctcag 4620

agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc ccgatctagt 4680

aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct 4740

atcgcgatt aaatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac 4800

gcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4860

atcatcgcaa gaccggcaac aggattcaat cttaagaaac ttatttgcca aatgtttgaa 4920

cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 4980

atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt gatgtggacg 5040

ccgggcccga tcatattgtc gctcaggatc gtggcggtgt gcttgtcggc cgttgctgtc 5100

gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtaggc gaagaactcc 5160

agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac gattccgaag 5220

cccaaccttt catagaaggc ggcgggtggaa tcgaaatctc gtgatggcag gttgggcgtc 5280

gcttggtcgg tcatttcgaa cccagagtc ccgtcagaa gaactcgtca agaaggcgat 5340

agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 5400

ccattcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg tcctgatagc 5460

ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca tttccacca 5520

tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc 5580

gcgccttgag cctggcgaa agttcggctg gcgcgagccc ctgatgctct tcgtccagat 5640

catcctgac gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 5700

cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc attgcatcag 5760

ccatgatgga tacttttctc gcaggagcaa ggtgagatga caggagatcc tgccccggca 5820

cttcgcccga tagcagccag tcccttcccg cttcagtgac aacgtcgagc acagctgcgc 5880

aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc agttcattca 5940

gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga 6000

acacggcggc atcagagcag ccgattgtct gttgtgccc gtcatagccg aatagcctct 6060

ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgac 6120

cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 6180

gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg 6240

accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa 6300

actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca gttccaaacg 6360

taaaacggct tgtcccgct catcgccggg ggtcataacg tgactccctt aattctccgc 6420

tcatgatcag attgtcggtt cccgccttca gtttaaacta tcagtgtttg acaggatata 6480

ttggcgggta aacctagag aaaagagcgt ttattagaat aatcgatat taaaagggc 6540

gtgaaaagg tttatccgtt gtccatttgt atgtgcatgc caaccacagg gttccccaga 6600

tctggcgccg gccagcgaga cgagcaagat tggccgcgc cggaaacgat ccgacagcg 6660

gcccagcaca ggtgcgagg caaattgcac caacgcatac agcgccagca gaatgccata 6720

gtgggcgggt acgtcggtt agtgaaccag atcgcgagg agggccggca gcaccggcat 6780

aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga tcaggggtat 6840

gttgggttt acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggatt 6900

ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg 6960

acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggctggc 7020

gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca gccggcgctt 7080

tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc catgctggcg 7140

gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt tctgatcggg 7200

aatgcccga gcttcaggca ggcgctgtc gcctaccgcg atggcgcgcg catccatgcc 7260

ggcacgcgac cgggocgacc gcagatggaa acggccgacg cgcagcttcg cttcctctgc 7320

gaggcggggt tttcgcccg ggacgccgtc aatgcgctga tgacaatcag ctacttcact 7380

gttggggccg tgcttgagga gcaggccggc gacagcgatg cggcgagcg cggcggcacc 7440
gttgaacagg ctccgctctc gccgctgttg cgggccgcga tagacgcctt cgacgaagcc 7500
gggccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg 7560
aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc aggaccgctg 7620
ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct ccccctttcc 7680
accgcgtcag acgcccgtag cagcccgtc cgggcttttt catgccctgc cctagcgtcc 7740
aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttcctcgc 7800
tcaactgactc gctgcgctcg gtcgttcggc tgccggcgagc ggtatcagct cactcaaagg 7860
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 7920
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 7980
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 8040
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttcga 8100
ccctgccgt taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgcttttcc 8160
gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc atcctttttc 8220
gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttggtg tatccaacgg 8280
cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc cttcttcaact 8340
gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg ctggccggct 8400
accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag 8460
ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag 8520
gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa 8580
atcacggggc tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa tggcgacctg 8640

ggcgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgac ggcgcggttc 8700

ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc 8760

aaggatcatga tgggcgtggc ccgcccggag gcagagccat gactttttta gccgctaaaa 8820

cggcgggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga 8880

cttcgaggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcctttgcga 8940

cgctcaccgg gctgggtgcc ctgcgcgtg ggctggcggc cgtctatggc cctgcaaacg 9000

cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcg ccgcccgcgt tgtggatacc 9060

tcgcgaaaaa cttggccctc actgacagat gaggggcgga cgttgacact tgaggggccc 9120

actcaccgg cgcgcgcttg acagatgagg ggcaggctcg atttcggccg gcgacgtgga 9180

gctggccagc ctgcgaaac gccgaaaacg cctgatttta cgcgagtttc ccacagatga 9240

tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg 9300

acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgagggggcg 9360

acctattgac atttgagggg ctgtccacag gcagaaaac cagcatttgc aagggtttcc 9420

gccggttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa 9480

accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg aaggggggtg 9540

cccccccttc tcgaaccctc ccggcccgt aacgcgggc tcccatcccc ccaggggctg 9600

cgccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt 9660

gccgggatcg gggcagtaac gggatggcg atcagcccga gcgcgacgcc cggaagcatt 9720

gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc 9780

ggcctgggtg ggggcctgcc cttcacttcg gccgtcgggg cattcacgga cttcatggcg 9840

gggcgggcaa tttttacctt gggcattctt ggcatagtgg tcgcggtgc cgtgctcgtg 9900

ttcgggggtg cgataaaccc agcgaaccat ttgaggtgat aggtaagatt ataccgaggt 9960
atgaaaacga gaattggacc ttacagaat tactctatga agcgccatat ttaaaaagct 10020
accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat attgacaata 10080
ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg 10140
caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca taaaaacttg 10200
catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt ctatcataat 10260
tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgccc gatgactttg 10320
tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccaggtgc 10380
tgccctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct gattacgtgc 10440
agctttccct tcaggcggga ttcatacagc ggcagccat ccgtcatcca taccaccag 10500
tcaaagggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg ttcaccgaat 10560
acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca gcgctggcgc 10620
gatttagccc cgacatagcc ccactgttcg tccatttccg cgagacgat gacgtcactg 10680
cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga cgtaaaatcg 10740
tgttgaggcc aacgcccata atgcgggctg ttgccggca tccaacgcca ttcattggcca 10800
tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac tgcagttgcc 10860
atgttttacg gcagtgagag cagagatagc gctgatgtcc ggcggtgctt ttgccgttac 10920
gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10980
cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cataattgtg 11040
gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac aactttgaaa 11100
aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct 11160

tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 11220

aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11280

aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag aaaatgaaaa 11340

cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga 11400

aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc tgcactttga 11460

acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgctcc tttgctcgga 11520

agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcacag 11580

gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt 11640

agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga 11700

agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc 11760

cgaagaggaa cttgtctttt ccacggcga cctgggagac agcaacatct ttgtgaaaga 11820

tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtggatatga 11880

cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt atgtcgagct 11940

attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact 12000

ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 12060

ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc aagtatttgg 12120

gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac gagaaggacg 12180

gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12240

caccaggcgg gtcaaacag gaataagggc acattgcccc ggcgtgagtc ggggcaatcc 12300

cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12360

acgcgggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc atgcgtgcgc 12420

cccgcgaaac cttccagtcc gtcggctcga tgggtccagca agctacggcc aagatcgagc 12480
gcgacagcgt gcaactggct cccctgccc tgcccgcgcc atcggccgcc gtggagcgtt 12540
cgcgtcgtct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc gacacgcgag 12600
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg 12660
aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt 12720
ccttgttcga tattgcgcg tggccggaca cgatgcgagc gatgccaac gacacggccc 12780
gctctgcctt gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg caaaacaagg 12840
tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag ctgcgggccc 12900
acgatgacga actggtgtgg cagcagggtg tggagtacgc gaagcgcacc cctatcggcg 12960
agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggctg atcaatggcc 13020
ggtattacac gaaggccgag gaatgcctgt cgcgccctaca ggcgacggcg atgggcttca 13080
cgtccgaccg cgttgggcac ctggaatcgg tgctgctgct gcaccgcttc cgcgtcctgg 13140
accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc gtcgtgctgt 13200
ttgctggcga cactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg 13260
cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgctc aagctggaaa 13320
ccttccgcct catgtgcgga tcggattcca ccgcgtgaa gaagtggcgc gagcaggctg 13380
gcgaagcctg cgaagagttg cgaggcagcg gcctgggtgga acacgcctgg gtcaatgatg 13440
acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg ggttcagcag 13500
ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact tgcttcgctc 13560
agtatcgctc gggacgcacg gcgcgtcta cgaactgccg ataaacagag gattaaaatt 13620
gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc aggatttccg 13680

cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg tttacgagca 13740
cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg tggcattcgg 13800
cgcctacatc gacggcgaga tcattggggt gtcggtcttc aaacaggagg acggccccaa 13860
ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc gaggccgagg 13920
ggtcgccggt atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga tgatcgtccg 13980
acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac ttaatatctc 14040
gctattctgg agcttggtgt ttatttcggt ctaccgcctg cggggcgggg tcgcggcgac 14100
ggtaggcgct gtgcagccgc tgatggtcgt gttcatctct gccgctctgc taggtagccc 14160
gatacgattg atggcgggtcc tgggggctat ttgcggaact gcgggcgtgg cgctgttggt 14220
gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg cgggggcggt 14280
ttccatggcg ttcggaaccg tgctgacctg caagtggcaa cctcccgtgc ctctgctcac 14340
ctttaccgpc tggcaactgg cggccggagg acttctgctc gttccagtag ctttagtggt 14400
tgatccgcca atcccgatgc ctacaggaac caatgttctc ggctggcgt ggctcggcct 14460
gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac tcgaacctac 14520
agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc cggggatgca 14580
tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag caatggatag 14640
gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc ttctcagcg 14700
gctttatcca gcgatttcct attatgtcgg catagtctc aagatcgaca gcctgtcacg 14760
gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg agatgatatt 14820
tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 14880
atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14940

gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 15000
ctgtatcgag tgggtgatttt gtgccgagct gccggtcggg gagctggttg ctggctggtg 15060
gcaggatata ttgtgggtgta aacaaattga cgcttagaca acttaataac acattgcgga 15120
cgtttttaat gtactgggggt ggtttttctt ttcaccagtg agacgggcaa cagctgattg 15180
cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg ttgccccagc 15240
aggcgaaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat aaatcaaaag 15300
aatagcccga gatagggttg agtggtgttc cagtttggaa caagagtcca ctattaaaga 15360
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 15420
aaccatcacc caaatcaagt tttttggggg cgaggtgccg taaagcacta aatcggaacc 15480
ctaaagggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg 15540
aaggggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg ggaagggcga 15600
tcgggtgcggg cctcttcgct attacgccag ctggcgaaaag ggggatgtgc tgcaaggcga 15660
ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtga 15720
ttcgagctcg gtacccggg 15739

<210> 4

<211> 11611

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (227)..(227)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (318)..(318)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (526)..(526)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (8946)..(8946)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10028)..(10028)

<223> n is a, c, g, or t

<400> 4

agcttgcacg cctgcaggtc gagtggagat gtggagtggg cgcttacaca gtacacgagg	60
acttctagct agaaagaagg attacctcta aacaagtgta cctgtgcatt ctgggtaa	120
gactcatagg agagttgtaa aaaagtttcg gccggcgat tgggtgttac ggagcattca	180
ctaggcaacc atgcatgggt actattgtat accatcttag taggaantga tttcgagggt	240
tatacctacg atgaatgtgt gtctgttagg cttgagagtt caaggaagaa acatgcaatt	300
atctttgcga acccaggngc tgggtgacgga attttcatag tcaagctatc agagtaaaga	360
agaggagcat gtcaaagtac aattagagac aaatatatag tcgcgtggag ccaagagcgg	420
attcctcagt ctcgtaggtc tcttgacgac cgttgatctg cttgatctcg tctcccgaaa	480
atgaaaatag ctctgctaag ctattcttct cttcgccgga gcctgnaagg cgttactagg	540
ttgcagtcaa tgcattaatg cattgcagat gagctgtatc tggaagaggt aaacccgaaa	600

acgcgtttta ttcttggtga catggagcta ttaaactact agaaggcact ctttgctgct	660
tggacaaatg aacgtatctt atcgagatcc tgaacaccat ttgtctcaac tccggagctg	720
acatcgacac caacgatctt atatccagat tcgtcaagct gtttgatgat ttcagtaacg	780
ttaagtggat cgatccccg gtcggcatct actctattcc tttgccctcg gacgagtgct	840
ggggcgctcg tttccactat cggcgagtac ttctacacag ccatcggtcc agacggccgc	900
gcttctgcgg gcgatttggt tacgccccg agtccccggt cgggatcgga cgattgcgtc	960
gcatcgaccc tgcgcccagg ctgcatcatc gaaattgccg tcaaccaagc tctgatagag	1020
ttgggtcaaga ccaatgcgga gcatatacgc ccggagccgc ggcgatcctg caagctccgg	1080
atgcctccgc tcgaagtagc gcgtctgctg ctccatacaa gccaaaccag gcctccagaa	1140
gaagatggtg gcgacctgt attgggaatc cccgaacatc gcctcgctcc agtcaatgac	1200
cgctgttatg cggccattgt ccgtcaggac attgttgagg ccgaaatccg cgtgcacgag	1260
gtgccggact tcggggcagt cctcggccca aagcatcagc tcatcgagag cctgcgcgac	1320
ggacgcactg acggtgtcgt ccatcacagt ttgccagtga tacacatggg gatcagcaat	1380
cgcgcatatg aaatcacgcc atgtagtgtg ttgaccgatt ccttgccggtc cgaatgggcc	1440
gaacccgctc gtctggctaa gatcggccgc agcgatcgca tccatggcct ccgcgaccgg	1500
ctgcagaaca gcgggcagtt cggtttcagg caggctcttc aacgtgacac cctgtgcacg	1560
gcgggagatg caataggtca ggctctcgct gaattcccca atgtcaagca cttccggaat	1620
cgggagcgcg gccgatgcaa agtgccgata aacataacga tctttgtaga aaccatcggc	1680
gcagctatth acccgcagga catatccacg ccctcctaca tcgaagctga aagcacgaga	1740
ttcttcgccc tccgagagct gcatcaggtc ggagacgctg tcgaactttt cgatcagaaa	1800
cttctcgaca gacgtcgcg tgagttcagg catggtgatg tctgctcaag cggggtagct	1860

gttagtcaag ctgcatgaa gtgggaaagc tcgaactgaa aggttcaaag gaataagggg 1920

tgggaaggat ggagtatgga tgtagcaaag tacttactta ggggaaataa aggttcttgg 1980

atgggaagat gaatatactg aagatgggaa aagaaagaga aaagaaaaga gcagctggtg 2040

gggagagcag gaaaatatgg caacaaatgt tggactgacg caacgacctt gtcaaccccg 2100

ccgacacacc gggcggacag acggggcaaa gctgcctacc agggactgag ggacctcagc 2160

aggtcgagtg cagagcaccg gatgggtcga ctgccagctt gtgttcccgg tctgcgccgc 2220

tggccagctc ctgagcggcc tttccggttt catacaccgg gcaaagcagg agaggcacga 2280

tatttggacg ccctacagat gccggatggg ccaattaggg agcttacgcg ccgggtactc 2340

gctctaccta cttcggagaa ggtactatct cgtgaatctt ttaccagatc ggaagcaatt 2400

ggacttctgt acctagggtta atggcatgct atttcgccga cggctataca cccctggctt 2460

cacattctcc ttcgcttact gccggtgatt cgatgaagct ccatattctc cgatgatgca 2520

atagattctt ggtcaacgag gggcacacca gcctttccac ttcggggcgg aggggcggcc 2580

ggtcccggat taataatcat ccactgcacc tcagagccgc cagagctgtc tggccagtgg 2640

cttattactc agcccttctc tctgcgtccg tccgtctctc cgcatgccag aaagagtcac 2700

cggtcactgt acagagctca cgagttcgtc acatTTTTct acaaattggtg gaggcggcgg 2760

attttaggct caagtcatga ccctctgggt cactccagaa tcagctaggt caacgaataa 2820

ggatgattct ataggaagat ccaggcaccg gtcaaccatg atctggacag atttgggagc 2880

tcggtataag ctctccacct atcttattct gtatagttaa ggcttaaagt ttatccagga 2940

gatgttgctg aagtcgattt gagtccactt cctcactggg agctatacga ctttgatggg 3000

cgttgtaggg gctgtattag gtctcgatca aacacaaata gaattaaatg gtactcgagt 3060

ccactgaagg tggcttctcc gtcttccgta gccgtgccga aatccttaca gcttggtgtg 3120

tgtgactttt ggttacgccg tctgactttt gtgggtgagct aactagagat catgctatat 3180
ctcctgattt aatacaatgc tcatcataac attccacctg gaactgctag caacgtttga 3240
cttgcatgtt gcaacgccct ttgcagagct atcggatgat caatagtgcc acgttctaaa 3300
ttcaaccaac gcagggtgcc caagccttcg acatccggat gtatttcgaa aacctcatgg 3360
cgattgcagt cctcagattc atgttcattc caatgctcat tggatgaataa aagggtcaca 3420
gggaataagt tcaaactcga gatacttgag aatattgaaa gccaaaggac cctctatgct 3480
ccaagctaga gtctcagcct ggaaagcaaa tccaaatgaa gctatgctac ctccaattcc 3540
tcatcatctt atctataata cagagtcgaa gaatatctc ttgacaccgc tccgtcctcc 3600
gacttcaata aggagcttac tcctccttga caccaccctt ccagttcttc tcggcgcttct 3660
ggagggaggc cttgtcggtc ttgggctggc cctggctgag aaagctgttg gcagccttaa 3720
agggacgctg gaggtcacca gtcgctggct tccgaagac gtggatctta accagattcg 3780
aaagcgcctt cagcggatga tcgactggat cagaagagcg ttggtgtact tgaagtacag 3840
atgcatgacg gccatcatgc caacgcccat gaactggctc ttaatgagct ggcggaactg 3900
gcccttatcg tactccatgt tggtagttgt gacaggacga ggctcctcgc cgcttccaag 3960
cggagcaggc tcgacgtatt tcagtgtcga aagatctgat caagagacag gatgaggatc 4020
gtttcgcatt attgaacaag atggattgca cgcagggttct ccggccgctt ggggtggagag 4080
gctattcggc tatgactggg cacaacagac aatcggctgc tctgatgccg ccgtgttccg 4140
gctgtcagcg cagggggcgcc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa 4200
tgaactgcag gacgaggcag cgcggctatc gtggctggcc acgacgggag ttccttgccg 4260
agctgtgctc gacgttgtca ctgaagcggg aagggaactg ctgctattgg gcgaagtgcc 4320
ggggcaggat ctctgtcat ctcaccttgc tcctgccgag aaagtatcca tcatggctga 4380

tgcaatgcgg cggtgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa 4440
acatcgcatc gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct 4500
ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcat 4560
gcccgcgggc gaggatctcg tcgtgacca tggcgatgcc tgcttgccga atatcatggt 4620
ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta 4680
tcaggacata gcgttggtta cccgtgatat tgctgaagag cttggcggcg aatgggctga 4740
ccgcttcctc gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg ccttctatcg 4800
ccttcttgac gagttcttct gagcgggact ctggggttcg aaatgaccga ccaagcgacg 4860
cccaacctgc catcacgaga tttcgattcc accgcgcct tctatgaaag gttgggcttc 4920
ggaatcgttt tccgggacgc cggtcggatg atcctccagc gcggggatct catgctggag 4980
ttcttcgccc accccgggct cgatccctc gcgagttggt tcagctgctg cctgaggctg 5040
gacgacctcg cggagttcta ccggcagtgc aaatccgtcg gcatccagga aaccagcagc 5100
ggctatccgc gcatccatgc ccccgaaactg caggagtggg gaggcacgat ggccgctttg 5160
gtccggatct ttgtgaagga accttacttc tgtggtgtga cataattgga caaactacct 5220
acagagattt aaagctctaa ggtaaatata aaatttttaa gtgtataatg tgttaaacta 5280
ctgattctaa ttgtttgtgt atttttagatt ccaacctatg gaactgatga atgggagcag 5340
tggtggaatg cttttaatga ggaaaacctg ttttgctcag aagaaatgcc atctagtgat 5400
gatgaggcta ctgctgactc tcaacattct actcctcaa aaaagaagag aaaggtagaa 5460
gacccaagg actttccttc agaattgcta agttttttga gtcattgctgt gtttagtaat 5520
agaactcttg cttgctttgc tatttacacc acaaaggaaa aagctgcact gctatacaag 5580
aaaattatgg aaaaatatc tgtaaccttt ataagtaggc ataacagtta taatcataac 5640

atactgtttt ttcttactcc acacagggcat agagtgtctg ctattaataa ctatgctcaa 5700

aaattgtgta ccttttagctt ttttaatttgt aaaggggtta ataaggaata tttgatgtat 5760

agtgcccttga ctagagatca taatcagcca taccacattt gtagagggtt tacttgcttt 5820

aaaaaacctc ccacacctcc ccctgaacct gaaacataaa atgaatgcaa ttgttggtgt 5880

taacttgttt attgcagctt ataatggta caaataaagc aatagcatca caaatctcac 5940

aaataaagca tttttttcac tgcattctag ttgtgggttg tccaaactca tcaatgtatc 6000

ttatcatgtc tggatctgac ggggtgcgcat gatcgtgctc ctgctggtga ggacccggct 6060

aggctggcgg ggttgcctta ctggtttagca gaatgaatca ccgatacgcg agcgaacgtg 6120

aagcgactgc tgctgcaaaa cgtctgcgac ctgagcaaca acatgaatgg tcttcgggtt 6180

ccgtgtttcg taaagtctgg aaacgcggaa gtcagcgctc ttccgcttcc tcgctcactg 6240

actcgctgcg ctcggtcggt cggctgcggc gagcggatc agctcactca aaggcggtaa 6300

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 6360

aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 6420

gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 6480

gacaggacta taaagatacc aggcgtttcc ccctggaagc tcctcgtgc gctctcctgt 6540

tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 6600

ttctcatagc tcacgctgta ggtatctcag ttcgggtgtag gtcgttcgct ccaagctggg 6660

ctgtgtgcac gaaccccccg ttcagccga ccgctgcgcc ttatccggta actatcgtct 6720

tgagtccaac ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat 6780

tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtggc ctaactacgg 6840

ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 6900

aagagttggg agctcttgat ccggcaaaca aaccaccgct ggtagcgggtg gtttttttgt 6960

ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 7020

tacgggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt 7080

atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 7140

aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 7200

ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg ttagataaac 7260

tacgatacgg gagggcttac catctggccc cagtgcctgca atgataccgc gagaccacg 7320

ctcaccgggt ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag 7380

tggtcctgca actttatccg cctccatcca gtctattaat tggtgccggg aagctagagt 7440

aagtagttcg ccagttaata gtttgcgcaa cgttggtgcc attgctgcag gcatcgtggg 7500

gtcacgctcg tcgtttggta tggcttcatt cagctccggg tcccaacgat caaggcgagt 7560

tacatgatcc cccatgttgt gcaaaaaagc ggtagctcc ttcggtcctc cgatcgttgt 7620

cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct 7680

tactgtcatg ccatccgtaa gatgcttttc tgtgactggg gagtactcaa ccaagtcatt 7740

ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaacac gggataatac 7800

cgcgccacat agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa 7860

actctcaagg atcttaccgc tggtgagatc cagttcgatg taaccactc gtgcacccaa 7920

ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 7980

aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct 8040

ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 8100

atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 8160

tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag 8220

gccctttcgt cttcaagaat tcgcgggccgc aattaaccct cadtaaagga tccctatagt 8280

gagtcgtatt atgcggccgc gaattctcat gtttgaccgc ttatcatcga taagctctgc 8340

tttttgttga cttccattgt tcattccacg gacaaaaaca gagaaaggaa acgacagagg 8400

ccaaaaagct cgctttcagc acctgtcgtt tcctttcttt tcagagggta ttttaaataa 8460

aaacattaag ttatgacgaa gaagaacgga aacgccttaa accggaaaat tttcataaat 8520

agcgaaaacc cgcgaggtcg ccgccccgta acaaggcgga tcgccggaaa ggacccgcaa 8580

atgataataa ttatcaattg catactatcg acggcactgc tgccagataa caccaccggg 8640

gaaacattcc atcatgatgg ccgtgcggac ataggaagcc agttcatcca tcgctttctt 8700

gtctgctgcc atttgctttg tgacatocag cgccgcacat tcagcagcgt ttttcagcgc 8760

gttttcgatc aacgtttcaa tgttgggtatc aacaccaggt ttaactttga acttatcggc 8820

actgacgggt accttgttct gcgctggctc atcacgcagg ataccaaggc tgatgttgta 8880

gatattggtc accggctgag ggttttcgat tgccgctgcg tggatagcac catttgcgat 8940

caggcngtcc ttgatgaatg aactccatt gcgaataagt tcgaaggaga cgggtgtcacg 9000

aatgcgctgg tccagctcgg tcgattgcct tttgtgcagc agaggtatca atctcaacgc 9060

caaggctcat cgaagcgcaa tattgctgct caccaaacg cgtattgacc aggtgttcaa 9120

cggcaaattt ctgcccttct gatgtcagaa aggcaaagtg attttctttc tggatttcag 9180

ttgctgtgtg tcggtttcag caaaaccaag ctgcgcgaat tcggctgtgc agatttagaa 9240

ggcagatcac cagacagcaa cggccaacgg aaaacagcgc atacagaaca tccgtcgccg 9300

cgccgacaac gtgataattt ttatgacca tgatttattt ctttttagac gtgagcctgt 9360

cgcacagcaa agccgccgaa agttcctcga agctagcttc agacgtgtct agatacgtct 9420

gctttttgtt gacttccatt gttcattcca cggacaaaaa cagagaaagg aaacgacaga 9480

ggccaaaaag ctcgctttca gcacctgtcg tttcctttct tttcagaggg tatttttaaat 9540

aaaaacatta agttatgacg aagaagaacg gaaacgcctt aaaccggaaa attttcataa 9600

atagcgaaaa cccgcgaggt cgccgccccg taacaaggcg gatcgccgga aaggaccgcg 9660

aatgataat aattatcaat tgcatactat cgacggcact gctgccagat aacaccaccg 9720

gggaaacatt ccatcatgat ggccgtgcgg acataggaag ccagttcatc catcgctttc 9780

ttgtctgctg ccatttgctt tgtgacatcc agcgccgcac attcagcagc gtttttcagc 9840

gcgtttttcga tcaacgtttc aatgttggtg tcaacaccag gtttaacttt gaacttatcg 9900

gcactgacgg ttaccttggt ctgcgctggc tcatcacgca ggataccaag gctgatgttg 9960

tagatattgg tcaccggctg agggttttcg attgccgctg cgtggatagc accatttgcg 10020

atcaggcngt cttgatgaa tgacactcca ttgcgaataa gttcgaagga gacggtgtca 10080

cgaatgcgct ggtccagctc ggtcgattgc cttttgtgca gcagaggtat caatctcaac 10140

gccaaggctc atcgaagcgc aatattgctg ctcacaaaaa cgcgtattga ccagggtgttc 10200

aacggcaa at ttctgccctt ctgatgtcag aaaggcaaag tgattttctt tctggtattc 10260

agttgctgtg tgtcgggtttc agcaaaacca agctcgcgca attcggctgt gcagatttag 10320

aaggcagatc accagacagc aacggccaac ggaaaacagc gcatacagaa catccgtcgc 10380

cgcgccgaca acgtgataat ttttatgacc catgatttat ttccttttag acgtgagcct 10440

gtcgcacagc aaagccgccc aaagtctctc gaccgatgcc cttgagagcc ttcaaccag 10500

tcagctcctt ccggtgggcg cggggcatga ctatcgtcgc cgcacttatg actgtcttct 10560

ttatcatgca actcgtagga caggtgccgg cagcgctctg ggtcattttc ggcgaggacc 10620

gctttcgctg gagcgcgacg atgacgggc tgctgcttgc ggtattcgga atcttgcacg 10680
ccctcgctca agccttcgtc actgggtccg ccaccaaacg ttctggcgag aagcaggcca 10740
ttatcgccgg catggcgggc gacgcgctgg gctacgtctt gctggcgctt gcgacgcgag 10800
gctggatggc cttccccatt atgattcttc tcgcttcgg cggcacggg atgcccgcgt 10860
tgcaggccat gctgtccagg caggtagatg acgaccatca gggacagctt caaggatcgc 10920
tcgcggctct taccagccta acttcgatca ttggaccgct gatcgtcacg gcgatttatg 10980
ccgcctcggc gagcacatgg aacgggttgg catggattgt aggcgccgcc ctataccttg 11040
tctgcctccc cgcgttgct cgcggtgcat ggagccgggc cacctcgacc tgaatggaag 11100
ccggcggcac ctcgctaacg gattcaccac tccaagaatt ggagccaatc aattcttgcg 11160
gagaactgtg aatgcgcaaa ccaacccttg gcagaacata tccatcgct cgcctctc 11220
cagcagccgc acgcggcgca tctcgggcag cgttgggtcc tgcagatccg gctgtggaat 11280
gtgtgtcagt tagggtgtgg aaagtcccca ggctccccag caggcagaag tatgcaaagc 11340
atgcatctca attagtcagc aaccaggtgt ggaaagtccc caggctcccc agcaggcaga 11400
agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgccct aactccgcc 11460
atcccccccc taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt 11520
tttatttatg cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga 11580
ggcttttttg gaggcctagg cttttgcaaa a 11611

<210> 5

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 5

cgatgtagga gggcgtggat a

21

<210> 6

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 6

gcttctgcgg gcgatttg t

21

<210> 7

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 7

tgagaatatc accggaattg

20

<210> 8

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 8

agctcgacat actgttcttc c

21

9

24

.> DNA

13> Artificial

<220>

<223> Primer

<400> 9

gtgaatggaa atcccatcgc tgtc

24

<210> 10

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 10

agtgggtact ctaaaggcca tacc

24

<210> 11

<211> 1771

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (166)..(1155)

<400> 11

ggcagagct tgcacgcaag tcagcgcgcg caagtcaaca cctgccgggc cacagcctca . 60

aataataaag agctcaagcg tttgtgcgcc tcgacgtggc cagtctgcac tgccttgaac 120

ccgcgagtct cccgccgcac tgactgcat agcacagcta gacga atg cag cta gca 177

Met Gln Leu Ala

gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag 225
 Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys
 5 10 15 20

gag aag gag aag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg 273
 Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp
 25 30 35

gcg acc cag tac tcg ctt ccg tca gaa gag tca gac gcg gcc cgc ccg 321
 Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp Ala Ala Arg Pro
 40 45 50

gga ctg aag aat gcc tac aag cca cca cct tcc gac aca aag ggc atc 369
 Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp Thr Lys Gly Ile
 55 60 65

aca atg gcg cta cgt gtc atc ggc tcc tgg gcc gca gtg ttc ctc cac 417
 Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala Val Phe Leu His
 70 75 80

gcc att ttt caa atc aag ctt ccg acc tcc ttg gac cag ctg cac tgg 465
 Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp Gln Leu His Trp
 85 90 95 100

ctg ccc gtg tca gat gcc aca gct cag ctg gtt agc ggc acg agc agc 513
 Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser Gly Thr Ser Ser
 105 110 115

ctg ctc gac atc gtc gta gta ttc ttt gtc ctg gag ttc ctg tac aca 561
 Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu Phe Leu Tyr Thr
 120 125 130

ggc ctt ttt atc acc acg cat gat gct atg cat ggc acc atc gcc atg 609
 Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly Thr Ile Ala Met
 135 140 145

aga aac agg cag ctt aat gac ttc ttg ggc aga gta tgc atc tcc ttg 657
 Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val Cys Ile Ser Leu
 150 155 160

31/357

tac gcc tgg ttt gat tac aac atg ctg cac cgc aag cat tgg gag cac	705
Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys His Trp Glu His	
165 170 175 180	
cac aac cac act ggc gag gtg ggc aag gac cct gac ttc cac agg gga	753
His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp Phe His Arg Gly	
185 190 195	
aac cct ggc att gtg ccc tgg ttt gcc agc ttc atg tcc agc tac atg	801
Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met	
200 205 210	
tcg atg tgg cag ttt gcg cgc ctc gca tgg tgg acg gtg gtc atg cag	849
Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr Val Val Met Gln	
215 220 225	
ctg ctg ggt gcg cca atg gcg aac ctg ctg gtg ttc atg gcg gcc gcg	897
Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala	
230 235 240	
ccc atc ctg tcc gcc ttc cgc ttg ttc tac ttt ggc acg tac atg ccc	945
Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Met Pro	
245 250 255 260	
cac aag cct gag cct ggc gcc gcg tca ggc tct tca cca gcc gtc atg	993
His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser Pro Ala Val Met	
265 270 275	
aac tgg tgg aag tcg cgc act agc cag gcg tcc gac ctg gtc agc ttt	1041
Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp Leu Val Ser Phe	
280 285 290	
ctg acc tgc tac cac ttc gac ctg cac tgg gag cac cac cgc tgg ccc	1089
Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His His Arg Trp Pro	
295 300 305	
ttc gcc ccc tgg tgg gag ctg ccc aac tgc cgc cgc ctg tct ggc cga	1137
Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg Leu Ser Gly Arg	
310 315 320	
ggg ctg gtt cct gcc tag ctggacacac tgcagtgggc cctgctgcca	1185
Gly Leu Val Pro Ala	

325

gctgggcatg caggttgtgg caggactggg tgaggtgaaa agctgcaggc gctgctgccg 1245
 gacacgctgc atgggctacc ctgtgtagct gccgccacta ggggagggggg tttgtagctg 1305
 tcgagcttgc cccatggatg aagctgtgta gtggtgcagg gagtacaccc acaggccaac 1365
 acccttgacg gagatgtctt gcgtcgggag gagtgttggg cagtgtagat gctatgattg 1425
 tatcttaatg ctgaagcctt taggggagcg acacttagtg ctgggcaggc aacgccctgc 1485
 aaggtgcagg cacaagctag gctggacgag gactcgggtg caggcagggtg aagaggtgcg 1545
 ggaggggtggg gccacacca ctgggcaaga ccatgctgca atgctggcgg tgtggcagtg 1605
 agagctgcgt gattaactgg gctatggatt gtttgagcag tctcacttat tctttgatat 1665
 agatactggg caggcaggtc aggagagtga gtatgaacaa gttgagaggt ggtgcgctgc 1725
 ccctgcgctt atgaagctgt aacaataaag tggttcaaaa aaaaaa 1771

<210> 12

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 12

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala
 1 5 10 15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val
 20 25 30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp
 35 40 45

33/357

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp
50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala
65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp
85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser
100 105 110

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr

210

215

220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser
260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp
275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His
290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg
305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala
325

<210> 13

<211> 1662

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (168) .. (1130)

<400> 13

cggggcaact caagaaattc aacagctgca agcgcgcccc agcctcacag cgccaagtga 60
 gctatcgacg tggttgtgag cgctcgacgt ggtccactga cgggcctgtg agcctctgcg 120
 ctccgtcctc tgccaaatct cgcgtcgggg cctgcctaag tcgaaga atg cac gtc 176
 Met His Val
 1
 gca tcg gca cta atg gtc gag cag aaa ggc agt gag gca gct gct tcc 224
 Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser
 5 10 15
 agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc 272
 Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser
 20 25 30 35
 gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct 320
 Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala Tyr Lys Pro
 40 45 50
 cca gca tct gac gcc aag ggc atc acg atg gcg ctg acc atc att ggc 368
 Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr Ile Ile Gly
 55 60 65
 acc tgg acc gca gtg ttt tta cac gca ata ttt caa atc agg cta ccg 416
 Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile Arg Leu Pro
 70 75 80
 aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa gcc aca gcc 464
 Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu Ala Thr Ala
 85 90 95
 cag ctt ttg ggc gga agc agc agc cta ctg cac atc gct gca gtc ttc 512
 Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala Ala Val Phe
 100 105 110 115
 att gta ctt gag ttc ctg tac act ggt cta ttc atc acc aca cat gac 560
 Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp
 120 125 130
 gca atg cat ggc acc ata gct ttg agg cac agg cag ctc aat gat ctc 608
 Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu

135	140	145	
ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac tac agc atg			656
Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Ser Met			
150	155	160	
ctg cat cgc aag cac tgg gag cac cac aac cat act ggc gaa gtg ggg			704
Leu His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly			
165	170	175	
aaa gac cct gac ttc cac aag gga aat ccc ggc ctt gtc ccc tgg ttc			752
Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe			
180	185	190	195
gcc agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg			800
Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu			
200	205	210	
gca tgg tgg gca gtg gtg atg caa atg ctg ggg gcg ccc atg gca aat			848
Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro Met Ala Asn			
215	220	225	
ctc cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc			896
Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu			
230	235	240	
ttc tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca			944
Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala			
245	250	255	
gca ggc tct cag gtg atg gcc tgg ttc agg gcc aag aca agt gag gca			992
Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr Ser Glu Ala			
260	265	270	275
tct gat gtg atg agt ttc ctg aca tgc tac cac ttt gac ctg cac tgg			1040
Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp			
280	285	290	
gag cac cac agg tgg ccc ttt gcc ccc tgg tgg cag ctg ccc cac tgc			1088
Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu Pro His Cys			
295	300	305	

cgc cgc ctg tcc ggg cgt ggc ctg gtg cct gcc ttg gca tga 1130
 Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala
 310 315 320

cctggtcctt ccgctggtga cccagcgtct gcacaagagt gtcattgtac. aggggtgctgc 1190

ggccagtggc agcgcagtgc actctcagcc tgtatggggc taccgctgtg ccactgagca 1250

ctgggcatgc cactgagcac tgggcgtgct actgagcaat gggcgtgcta ctgagcaatg 1310

ggcgtgctac tgacaatggg cgtgctactg gggctctggca gtggctagga tggagtttga 1370

tgcattcagt agcgggtggcc aacgtcatgt ggatgggtgga agtgctgagg ggtttaggca 1430

gccggcattt gagaggggcta agttataaat cgcattgctgc tcatgcgcac atatctgcac 1490

acagccaggg aaatcccttc gagagtgatt atgggacact tgtattgggtt tcgtgctatt 1550

gttttattca gcagcagtac ttagtgaggg tgagagcagg gtggtgagag tggagtgagt 1610

gagtatgaac ctggtcagcg aggtgaacag cctgtaatga atgactctgt ct 1662

<210> 14

<211> 320

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala
 1 5 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His
 20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala
 35 40 45

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr
50 55 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile
65 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu
85 90 95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala
100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr
115 120 125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu
130 135 140

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp
145 150 155 160

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly
165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val
180 185 190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe
195 200 205

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro

39/357

210

215

220

Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala
 225 230 235 240

Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro
 245 250 255

Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr
 260 265 270

Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp
 275 280 285

Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu
 290 295 300

Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala
 305 310 315 320

<210> 15

<211> 729

<212> DNA

<213> Agrobacterium aurantiacum

<220>

<221> CDS

<222> (1)..(729)

<400> 15

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg

48

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu

1

5

10

15

atc gtc tcg ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat	96
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His	
20 25 30	
gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca	144
Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala	
35 40 45	
aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg	192
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala	
50 55 60	
cat gac gcg atg cac ggg tcg gtg gtg ccg ggg cgt ccg cgc gcc aat	240
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn	
65 70 75 80	
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg	288
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp	
85 90 95	
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc	336
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr	
100 105 110	
gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc	384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala	
115 120 125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc	432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro	
130 135 140	
gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac	480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr	
145 150 155 160	
gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gcg tcg atc cag ctg ttc	528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe	
165 170 175	
gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg	576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro	

```
<210> 16
<211> 242
<212> PRT
<213> Agrobacterium aurantiacum
<400> 16
```

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp

225

230

235

240

Thr Ala

<210> 17

<211> 1631

<212> DNA

<213> Alcaligenes sp.

<220>

<221> CDS

<222> (99) .. (827)

<400> 17

ctgcaggccg ggcccgggtgg ccaatgggtcg caaccggcag gactggaaca ggacggcggg 60

ccgggtctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct 116

Met Ser Gly Arg Lys Pro

1

5

ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc 164

Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile

10

15

20

ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat 212

Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp

25

30

35

gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc 260

Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr

40

45

50

tgg ctg tgc gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg 308

Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly

55

60

65

70

tcc gtg gtg ccg ggg cgg ccg cgc gcc aat gcg gcg atc ggg caa ctg 356

Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu

75	80	85	
gcg ctg tgg ctc tat gcg ggg ttc tcg tgg ccc aag ctg atc gcc aag			404
Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys			
90	95	100	
cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc			452
His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe			
105	110	115	
ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat			500
Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr			
120	125	130	
ttc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat			548
Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro Val Ile Val Thr Thr Tyr			
135	140	145	150
gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc			596
Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val			
155	160	165	
ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg			644
Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu			
170	175	180	
ccc cac cgc ccg gga cat gac gat ttt ccc gac cgg cac aac gcg agg			692
Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg			
185	190	195	
tcg acc ggc atc ggc gac ccg ttg tca cta ctg acc tgc ttc cat ttc			740
Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe			
200	205	210	
ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg			788
Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp			
215	220	225	230
cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct			837
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala			
235	240		

cattgtcgtg gcgacagtcc tcgtgatgga gctgaccgcc tattccgtcc accgctggat 897
tatgcacggc cccctaggct ggggctggca caagtcccat cacgaagagc acgaccacgc 957
gttggagaag aacgacctct acggcgtcgt cttcgcggtg ctggcgacga tcctcttcac 1017
cgtggggcgcc tattggtggc cgggtgctgtg gtggatcgcc ctgggcatga cggcttatgg 1077

gttgatctat ttcatactgc acgacgggct tgtgcatcaa cgctggccgt ttcggtatat 1137
tccgcggcgg ggctatctcc gcaggctcta ccaagctcat cgcctgcacc acgcggtcga 1197
ggggcgggac cactgcgtca gcttcggctt catctatgcc ccacccgtgg acaagctgaa 1257
gcaggatctg aagcggtcgg gtgtcctgcg ccccaggac gagcgtccgt cgtgatctct 1317
gatcccgcg tggccgcatg aaatccgacg tgctgctggc aggggcccgc cttgccaacg 1377
gactgatcgc gctggcgatc cgcaaggcgc ggcccagact tcgctgctg ctgctggacc 1437
gtgcggcggg cgcctcggac gggcatactt ggtcctgcca cgacaccgat ttggcgccgc 1497
actggctgga ccgcctgaag ccgatcaggc gtggcgactg gcccgatcag gaggtgcggt 1557
tcccagacca ttcgcgaagg ctccggggcg gatatggctc gatcgacggg cgggggctga 1617

tgcgtgcggt gacc 1631

<210> 18

<211> 242

<212> PRT

<213> *Alcaligenes* sp.

<400> 18

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu
1 5 10 15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe

20

25

30

Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu
35 40 45

Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr
100 105 110

Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly
115 120 125

Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly
 225 230 235 240

Arg Ala

<210> 19

<211> 729

<212> DNA

<213> Paracoccus marcusii

<220>

<221> CDS

<222> (1)..(729)

<400> 19

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc aca agc ctg 48
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtg cat 96
 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
 20 25 30

gcg ctg tgg ttt ctg gac gcg gcg gcc cat ccc atc ctg gcg gtc gcg 144
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
 35 40 45

aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192
 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala

50

55

60

cat gac gcg atg cac ggg tcg gtc gtg ccg ggg cgt ccg cgc gcc aat 240
 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn

65

70

75

80

gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg 288
 Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp

85

90

95

cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc 336
 Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr

100

105

110

gac gac gac cca gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc 384
 Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala

115

120

125

cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc 432
 Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro

130

135

140

gtc atc gtg acg gtc tat gcg ctg atc ctg ggg gat cgc tgg atg tac 480
 Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr

145

150

155

160

gtg gtc ttc tgg ccg ttg ccg tcg atc ctg gcg tcg atc cag ctg ttc 528
 Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe

165

170

175

gtg ttc ggc act tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg 576
 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro

180

185

190

gac cgc cat aat gcg cgg tcg tcg cgg atc agc gac cct gtg tcg ctg 624
 Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu

195

200

205

ctg acc tgc ttt cat ttt ggc ggt tat cat cac gaa cac cac ctg cac 672
 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His

210

215

220

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720
 Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
 225 230 235 240

acc gca tga 729
 Thr Ala

<210> 20

<211> 242

<212> PRT

<213> Paracoccus marcusii

<400> 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
 20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
 35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
 50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
 65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
 85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
 100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

Thr Ala

<210> 21
<211> 1629
<212> DNA

<213> Synechocystis sp.

<220>

<221> CDS

<222> (1)..(1629)

<400> 21

atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta 48
Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
1 5 10 15

gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta 96
Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu
20 25 30

gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg 144
Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met
35 40 45

ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac 192
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His
50 55 60

gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag 240
Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln
65 70 75 80

tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg 288
Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly
85 90 95

ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt 336
Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys
100 105 110

gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa 384
Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125

ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt 432

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe	
130 135 140	
aat gct ccg ccc cag gct tta cta gat tta gcc ctg aac tat ggt tgg	480
Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp	
145 150 155 160	
gaa aac tta aaa tcc gtg ctg gcg atc gcc ggg tcg aaa acc aag gcg	528
Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala	
165 170 175	
ttg gat ttt atc cgc act atg atc ggc tcc ccg gaa gat gtg ctc aat	576
Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn	
180 185 190	
gaa tgg ttc gac agc gaa cgg gtt aaa gct cct tta gct aga cta tgt	624
Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys	
195 200 205	
tcg gaa att ggc gct ccc cca tcc caa aag ggt agt agc tcc ggc atg	672
Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met	
210 215 220	
atg atg gtg gcc atg cgg cat ttg gag gga att gcc aga cca aaa gga	720
Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly	
225 230 235 240	
ggc act gga gcc ctc aca gaa gcc ttg gtg aag tta gtg caa gcc caa	768
Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln	
245 250 255	
ggg gga aaa atc ctc act gac caa acc gtc aaa cgg gta ttg gtg gaa	816
Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu	
260 265 270	
aac aac cag gcg atc ggg gtg gag gta gct aac gga gaa cag tac cgg	864
Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg	
275 280 285	
gcc aaa aaa ggc gtg att tct aac atc gat gcc cgc cgt tta ttt ttg	912
Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu	

290	295	300	
caa ttg gtg gaa ccg ggg gcc cta gcc aag gtg aat caa aac cta ggg			960
Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly			
305	310	315	320
gaa cga ctg gaa cgg cgc act gtg aac aat aac gaa gcc att tta aaa			1008
Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys			
325	330	335	
atc gat tgt gcc ctc tcc ggt tta ccc cac ttc act gcc atg gcc ggg			1056
Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly			
340	345	350	
ccg gag gat cta acg gga act att ttg att gcc gac tcg gta cgc cat			1104
Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His			
355	360	365	
gtc gag gaa gcc cac gcc ctc att gcc ttg ggg caa att ccc gat gct			1152
Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala			
370	375	380	
aat ccg tct tta tat ttg gat att ccc act gta ttg gac ccc acc atg			1200
Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met			
385	390	395	400
gcc ccc cct ggg cag cac acc ctc tgg atc gaa ttt ttt gcc ccc tac			1248
Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr			
405	410	415	
cgc atc gcc ggg ttg gaa ggg aca ggg tta atg ggc aca ggt tgg acc			1296
Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr			
420	425	430	
gat gag tta aag gaa aaa gtg gcg gat cgg gtg att gat aaa tta acg			1344
Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr			
435	440	445	
gac tat gcc cct aac cta aaa tct ctg atc att ggt cgc cga gtg gaa			1392
Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu			
450	455	460	

agc ccc gcc gaa ctg gcc caa cgg ctg gga agt tac aac ggc aat gtc 1440
 Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
 465 470 475 480

tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta 1488
 Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
 485 490 495

ccg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca 1536
 Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
 500 505 510

ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga 1584
 Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
 515 520 525

aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa 1629
 Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
 530 535 540

<210> 22

<211> 542

<212> PRT

<213> Synechocystis sp.

<400> 22

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
 1 5 10 15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu
 20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met
 35 40 45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His
50 55 60

Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln
65 70 75 80

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly
85 90 95

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys
100 105 110

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe
130 135 140

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp
145 150 155 160

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala
165 170 175

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn
180 185 190

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys
195 200 205

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met

210

215

220

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly
225 230 235 240

Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln
245 250 255

Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu
260 265 270

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg
275 280 285

Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu
290 295 300

Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly
305 310 315 320

Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys
325 330 335

Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly
340 345 350

Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His
355 360 365

Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala
370 375 380

Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met
385 390 395 400

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr
405 410 415

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr
420 425 430

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr
435 440 445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu
450 455 460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
465 470 475 480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
485 490 495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
500 505 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
530 535 540

<211> 776

<212> DNA

<213> Bradyrhizobium sp.

<220>

<221> CDS

<222> (1) .. (774)

<400> 23

atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cgg cgc 48
Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg
1 5 10 15

gac gat gcg agg cag cgc cgc gtc ggt ctc acg ctg gcc gcg gtc atc 96
Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile
20 25 30

atc gcc gcc tgg ctg gtg ctg cat gtc ggt ctg atg ttc ttc tgg ccg 144
Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro
35 40 45

ctg acc ctt cac agc ctg ctg ccg gct ttg cct ctg gtg gtg ctg cag 192
Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln
50 55 60

acc tgg ctc tat gta ggc ctg ttc atc atc gcg cat gac tgc atg cac 240
Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His
65 70 75 80

ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga cag 288
Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln
85 90 95

ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc 336
Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val
100 105 110

gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat 384
Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp
115 120 125

59/357

ttc gac gag gtg ccg ccg cac ggc ttc tgg cac tgg ttc gcc agc ttt	432
Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe	
130 135 140	

ttc ctg cac tat ttc ggc tgg aag cag gtc gcg atc atc gca gcc gtc	480
Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val	
145 150 155 160	

tcg ctg gtt tat cag ctc gtc ttc gcc gtt ccc ttg cag aac atc ctg	528
Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu	
165 170 175	

ctg ttc tgg gcg ctg ccc ggg ctg ctg tcg gcg ctg cag ctg ttc acc	576
Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr	
180 185 190	

ttc ggc acc tat ctg ccg cac aag ccg gcc acg cag ccc ttc gcc gat	624
Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp	
195 200 205	

cgc cac aac gcg cgg acg agc gaa ttt ccc gcg tgg ctg tcg ctg ctg	672
Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu	
210 215 220	

acc tgc ttc cac ttc ggc ttt cat cac gag cat cat ctg cat ccc gat	720
Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp	
225 230 235 240	

gcg ccg tgg tgg cgg ctg ccg gag atc aag cgg cgg gcc ctg gaa agg	768
Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg	
245 250 255	

cgt gac ta	776
Arg Asp	

<210> 24

<211> 258

<212> PRT

<213> Bradyrhizobium sp.

<400> 24

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg
1 5 10 15

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile
20 25 30

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro
35 40 45

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln
50 55 60

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His
65 70 75 80

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln
85 90 95

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val
100 105 110

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp
115 120 125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe
130 135 140

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
145 150 155 160

61/357

Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu
 165 170 175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
 180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
 195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
 210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
 225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
 245 250 255

Arg Asp

<210> 25

<211> 777

<212> DNA

<213> Nostoc sp.

<220>

<221> CDS

<222> (1)..(777)

<400> 25

atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta
 Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu

1

5

10

15

48

ttg tca tcg aca atc aga gat gat aaa.aat att aat aag ggt ata ttt	96
Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe	
20 25 30	
att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta	144
Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu	
35 40 45	
ctc tca atā gat aca tcc ata att cat aag agc tta tta ggt ata gcc	192
Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala	
50 55 60	
atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat	240
Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His	
65 70 75 80	
gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat	288
Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn	
85 90 95	
ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa	336
Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys	
100 105 110	
gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat	384
Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp	
115 120 125	
tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg	432
Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp	
130 135 140	
tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga	480
Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly	
145 150 155 160	
tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa	528
Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu	
165 170 175	
aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta	576

63/357

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
 180 185 190

caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt 624
 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
 195 200 205

ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt 672
 Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
 210 215 220

tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac 720
 Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata 768
 Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

tct tta taa 777
 Ser Leu

<210> 26
 <211> 258
 <212> PRT
 <213> Nostoc sp.

<400> 26

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
 1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe
 20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu
 35 40 45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala
50 55 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
65 70 75 80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn
85 90 95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys
100 105 110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp
115 120 125

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp
130 135 140

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly
145 150 155 160

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu
165 170 175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
180 185 190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
195 200 205

65/357

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
 210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

Ser Leu

<210> 27

<211> 789

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(789)

<400> 27

ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa 48
 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
 1 5 10 15

tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta 96
 Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
 20 25 30

att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat 144
 Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
 35 40 45

tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa 192
 Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
 50 55 60

67/357

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac 768
 Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
 245 250 255

aat tca gta acc aat tcg taa 789
 Asn Ser Val Thr Asn Ser
 260

<210> 28

<211> 262

<212> PRT

<213> Nostoc punctiforme

<400> 28

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
 1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
 20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
 35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
 50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
 65 70 75 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
 85 90 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
245 250 255

Asn Ser Val Thr Asn Ser

260

<210> 29

<211> 762

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(762)

<400> 29

gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca	48
Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro	
1 5 10 15	

gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc	96
Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val	
20 25 30	

att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac	144
Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp	
35 40 45	

atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa	192
Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln	
50 55 60	

aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat	240
Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His	
65 70 75 80	

ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca	288
Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr	
85 90 95	

ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa	336
Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys	
100 105 110	

aaa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat 384
 Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
 115 120 125

ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt 432
 Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
 130 135 140

atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att 480
 Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
 145 150 155 160

tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act 528
 Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
 165 170 175

tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat 576
 Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
 180 185 190

ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag 624
 Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
 195 200 205

cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc 672
 Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
 210 215 220

acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat 720
 Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag 762
 Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
 245 250

<210> 30

<211> 253

<212> PRT

<213> Nostoc punctiforme

<400> 30

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro
1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp
35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr
85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys
100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
 165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
 180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
 195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
 210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
 245 250

<210> 31

<211> 1608

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (3) .. (971)

<400> 31

ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc 47
 Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile
 1 5 10 15

ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg 95

73/357

Gly	Pro	Pro	Pro	His	Leu	His	Arg	Ser	Phe	Ala	Ala	Thr	Thr	Met	Leu	
				20					25					30		
tcg	aag	ctg	cag	tca	atc	agc	gtc	aag	gcc	cgc	cgc	gtt	gaa	cta	gcc	143
Ser	Lys	Leu	Gln	Ser	Ile	Ser	Val	Lys	Ala	Arg	Arg	Val	Glu	Leu	Ala	
			35					40					45			
cgc	gac	atc	acg	cgg	ccc	aaa	gtc	tgc	ctg	cat	gct	cag	cgg	tgc	tcg	191
Arg	Asp	Ile	Thr	Arg	Pro	Lys	Val	Cys	Leu	His	Ala	Gln	Arg	Cys	Ser	
			50					55					60			
tta	gtt	cgg	ctg	cga	gtg	gca	gca	cca	cag	aca	gag	gag	gcg	ctg	gga	239
Leu	Val	Arg	Leu	Arg	Val	Ala	Ala	Pro	Gln	Thr	Glu	Glu	Ala	Leu	Gly	
			65					70					75			
acc	gtg	cag	gct	gcc	ggc	gcg	ggc	gat	gag	cac	agc	gcc	gat	gta	gca	287
Thr	Val	Gln	Ala	Ala	Gly	Ala	Gly	Asp	Glu	His	Ser	Ala	Asp	Val	Ala	
80					85				90					95		
ctc	cag	cag	ctt	gac	cgg	gct	atc	gca	gag	cgt	cgt	gcc	cgg	cgc	aaa	335
Leu	Gln	Gln	Leu	Asp	Arg	Ala	Ile	Ala	Glu	Arg	Arg	Ala	Arg	Arg	Lys	
				100					105					110		
cgg	gag	cag	ctg	tca	tac	cag	gct	gcc	gcc	att	gca	gca	tca	att	ggc	383
Arg	Glu	Gln	Leu	Ser	Tyr	Gln	Ala	Ala	Ala	Ile	Ala	Ala	Ser	Ile	Gly	
				115					120					125		
gtg	tca	ggc	att	gcc	atc	ttc	gcc	acc	tac	ctg	aga	ttt	gcc	atg	cac	431
Val	Ser	Gly	Ile	Ala	Ile	Phe	Ala	Thr	Tyr	Leu	Arg	Phe	Ala	Met	His	
			130					135					140			
atg	acc	gtg	ggc	ggc	gca	gtg	cca	tgg	ggt	gaa	gtg	gct	ggc	act	ctc	479
Met	Thr	Val	Gly	Gly	Ala	Val	Pro	Trp	Gly	Glu	Val	Ala	Gly	Thr	Leu	
			145					150					155			
ctc	ttg	gtg	gtt	ggt	ggc	gcg	ctc	ggc	atg	gag	atg	tat	gcc	cgc	tat	527
Leu	Leu	Val	Val	Gly	Gly	Ala	Leu	Gly	Met	Glu	Met	Tyr	Ala	Arg	Tyr	
160					165				170					175		
gca	cac	aaa	gcc	atc	tgg	cat	gag	tcg	cct	ctg	ggc	tgg	ctg	ctg	cac	575
Ala	His	Lys	Ala	Ile	Trp	His	Glu	Ser	Pro	Leu	Gly	Trp	Leu	Leu	His	
				180					185					190		

aag agc cac cac aca cct cgc act gga ccc ttt gaa gcc aac gac ttg 623
 Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu
 195 200 205

ttt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt acc ttt ggc 671
 Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly
 210 215 220

ttc tgg ctg ccc aac gtc ctg ggg gcg gcc tgc ttt gga gcg ggg ctg 719
 Phe Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu
 225 230 235

ggc atc acg cta tac ggc atg gca tat atg ttt gta cac gat ggc ctg 767
 Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu
 240 245 250 255

gtg cac agg cgc ttt ccc acc ggg ccc atc gct ggc ctg ccc tac atg 815
 Val His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met
 260 265 270

aag cgc ctg aca gtg gcc cac cag cta cac cac agc ggc aag tac ggt 863
 Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly
 275 280 285

ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att 911
 Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile
 290 295 300

cca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg 959
 Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp
 305 310 315

tcc aag cgg tag ggtgcggaac caggcagcgt ggtttcacac ctcatgcctg 1011
 Ser Lys Arg
 320

tgataagggtg tggctagagc gatgcgtgtg agacgggtat gtcacgggtcg actgggtctga 1071

tggccaatgg catcgcccat gtctgggtcat cacgggctgg ttgcctgggt gaaggtgatg 1131

cacatcatca tgtgcggttg gaggggctgg cacagtgtgg gctgaactgg agcagttgtc 1191

caggctggcg ttgaatcagt gagggtttgt gattggcggg tgtgaagcaa tgactccgcc 1251
catattctat ttgtgggagc tgagatgatg gcatgcttgg gatgtgcatg gatcatggta 1311
gtgcagcaaa ctatattcac ctagggctgt tggtaggata aggtgaggcc ttgcacattg 1371
catgatgtac tcgtcatggg gtgttgggtga gaggatggat gtggatggat gtgtattctc 1431
agacgtagac cttgactgga ggcttgatcg agagagtggg ccgtattctt tgagagggga 1491
ggctcgtgcc agaaatgggtg agtggatgac tgtgacgctg tacattgcag gcaggtgaga 1551
tgcactgtct cgattgtaaa atacattcag atgcaaaaaa aaaaaaaaaa aaaaaa 1608

<210> 32

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 32

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly
1 5 10 15

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser
20 25 30

Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg
35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu
50 55 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr
65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu
85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg
100 105 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val
115 120 125

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met
130 135 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu
145 150 155 160

Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala
165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys
180 185 190

Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe
195 200 205

Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe
210 215 220

Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly
225 230 235 240

Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val

77/357

245

250

255

His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys
260 265 270

Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly
275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro
290 295 300

Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser
305 310 315 320

Lys Arg

<210> 33

<211> 528

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)..(528)

<400> 33

atg ttg tgg att tgg aat gcc ctg atc gtt ttc gtt acc gtg att ggc 48
Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
1 5 10 15

atg gaa gtg att gct gca ctg gca cac aaa tac atc atg cac ggc tgg 96
Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
20 25 30

78/357

ggt tgg gga tgg cat ctt tca cat cat gaa ccg cgt aaa ggt gcg ttt 144
 Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
 35 40 45

gaa gtt aac gat ctt tat gcc gtg gtt ttt gct gca tta tcg atc ctg 192
 Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu
 50 55 60

ctg att tat ctg ggc agt aca gga atg tgg ccg ctc cag tgg att ggc 240
 Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly
 65 70 75 80

gca ggt atg acg gcg tat gga tta ctc tat ttt atg gtg cac gac ggg 288
 Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly
 85 90 95

ctg gtg cat caa cgt tgg cca ttc cgc tat att cca cgc aag ggc tac 336
 Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr
 100 105 110

ctc aaa cgg ttg tat atg gcg cac cgt atg cat cac gcc gtc agg ggc 384
 Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
 115 120 125

aaa gaa ggt tgt gtt tct ttt ggc ttc ctc tat gcg ccg ccc ctg tca 432
 Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser
 130 135 140

aaa ctt cag gcg acg ctc cgg gaa aga cat ggc gct aga gcg ggc gct 480
 Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala
 145 150 155 160

gcc aga gat gcg cag ggc ggg gag gat gag ccc gca tcc ggg aag taa 528
 Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys
 165 170 175

<210> 34

<211> 175

<212> PRT

<213> Erwinia uredovora

<400> 34

Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
1 5 10 15

Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
20 25 30

Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
35 40 45

Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu
50 55 60

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly
65 70 75 80

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly
85 90 95

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr
100 105 110

Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
115 120 125

Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser
130 135 140

Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala
145 150 155 160

Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys
165 170 175

<210> 35

<211> 1520

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 35

ctcgagtacc gaggcggaac ggcaggaatg tttccctctc ttttagaggg caattcttta 60
tccaatgtca tgttgatgct agatatttct gtctcttata ataaggcgaa taccattttt 120
tgaattgaag ttgagataaa aaaaaagggg gcccaatttg tcaacgcaa agagtcaagc 180
tttttctttg gctttagccg aacaatctaa gacttattgt ttttgaagat atttgacctt 240
ttctagatat tccttcaagt aaagcttttt tcgagttttt tttttttttc tttgtgaagg 300
atttattgtt attggtatcc attttttatt ggaagacaag ataagttaat attgattttg 360
cttaaagatt aaaaggaaat cagaaaacga caataaaaaa tgtaacggac aaactatggt 420
gtcgattata agtctaaatc cttaaaaaat gacaacgagt tgctttcctc tgaaaacaat 480
tcttttgtct ttgcaagaaa ggtttctttt ttgtttgctt gcattactta aacatcaaatt 540
caaatgaaag gaataaagca gatttgaggg cgaataagga ttttctgggc aacaagatgt 600
gagtgcaccc taaggaacta aatgccattc atttgtttta aaacgacatc aaagattgat 660
gatcaacagg attgagagag agaaaaagaa ctcgtgtcat ttatttctgt tgactgaaat 720
tttatattta gaaaaaatgt caaatctata gctttagcta tattacataa catttgaaat 780
aataataata aaaaaagaca cattagagac acttttcaaa ctctaaataa ctgtctataa 840

acacaaagaa aacaaagacc tctataacaa cttattagat ttttctcgta cttttgtcta 900
aagatgatgt attcttggtta tcccacactt ctttcatttg ttcttgatgc tactaaatat 960
acaaaatttc ttttttgcaa gagatattat tccaaaaatt ttcaaaaaga aatttttttc 1020
acaatagcag ttgatcgtgt aacccaaaga gggtctttgt tatttttgac ttccgctttg 1080
cggatgatgca tattcaaagt aatatatgga ataaacaacg tgtttaagca tgaaagaaag 1140
gaaacaaagg ccgctttgaa caaatgcata atatttcaga caaaaatgat ctaaagcaag 1200
cagtaaatca aacaagaaac attgctgatt cgcgttagaa aacgataaaa gtctaataag 1260
ccactaagta tacttcaatg aactttttgt atgcttatgg tccaatcaga ccaataattt 1320
gtgaccattc ctgagggtggc tttggtgatg cggaaacaga aaaaaatttt ctcaccaatc 1380
gatttaaaaa acaatttctg ctttgaacca aaactttttt tttctcttta atcattaact 1440
ttatcaagta tgtacctacc ctcaaagtcc tcactcaagc acaattatgc taacattggt 1500
ccaccttctc tttagaaatg 1520

<210> 36

<211> 16245

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 36

ccgggctggt	tgccctcgcc	gctgggctgg	cggccgtcta	tggccctgca	aacgcgccag	60
aaacgccgtc	gaagccgtgt	gcgagacacc	gcggccgccg	gcgttggtga	tacctgcggg	120
aaaacttggc	cctcactgac	agatgagggg	cggacgttga	cacttgaggg	gccgactcac	180
ccggcgccgc	gttgacagat	gaggggcagg	ctcgatttcg	gccggcgacg	tggagctggc	240
cagcctcgca	aatcggcgaa	aacgcctgat	tttacgcgag	tttcccacag	atgatgtgga	300
caagcctggg	gataagtgcc	ctgcggtatt	gacacttgag	gggcgcgact	actgacagat	360
gaggggcgcg	atccttgaca	cttgaggggc	agagtgctga	cagatgaggg	gcgcacctat	420
tgacatttga	ggggctgtcc	acaggcagaa	aatccagcat	ttgcaagggg	ttccgccogt	480
ttttcggcca	ccgctaacct	gtcttttaac	ctgcttttaa	accaatattt	ataaaccttg	540
tttttaacca	gggctgcgcc	ctgtgcgcgt	gaccgcgcac	gccgaagggg	ggtgcccccc	600
cttctcgaac	cctcccggcc	cgctaacgcg	ggcctcccat	ccccccaggg	gctgcgcccc	660
tcggccgcga	acggcctcac	cccaaaaatg	gcagcgctgg	cagtccttgc	cattgcgggg	720
atcggggcag	taacgggatg	ggcgatcagc	ccgagcgcga	cgcccggaag	cattgacgtg	780
ccgcaggtgc	tggcatcgac	attcagcgac	caggtgccgg	gcagtgaggg	cggcgggcctg	840
ggtggcggcc	tgcccttcac	ttcggccgtc	ggggcattca	cggacttcat	ggcgggggccg	900
gcaattttta	ccttgggcat	tcttggcata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa	accagcgaa	ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttataaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tctttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tataatctata gaatgggcaa agcataaaaaa cttgcatgga 1260

ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgetgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agtttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg gggtgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220

aattagcttc ttggggatatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc tcaaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgct ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtsagtggc tttattgatc ttggggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgct cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccc cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac caccgcgaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta caccgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctg 4500

tgcatcgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg caccggcgcg tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctgggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttattht cgggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtccctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcggga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgccctggcaa ctggcgggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttgggtccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacgggttaag 5820
cgagaaaatga ataagaaggc tgataattcg gatctctcgc agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgccgtccc gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300

aatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440

tcgaaatata tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa ggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaata 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggg catgacttga gcctaaaata cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggcg 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccgt gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttggtacat ccatactcca tccttcccat cccttattcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctaccccgc 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgccccga aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgt gttctgcagc cggtcgcgga 9300
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360
accgcaagga atcggatcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggc cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtag acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgccccca gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agtccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtcaa gcagcaaaga 10140

gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggc acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt aatctataca 10800

atgctccata gactcacatt gatattgtcg aagatttcga tgctgactta gtagagcaac 10860

tacaaaagtt agcagagaag catgatttct taatctttga agaccgcaag tttgcagata 10920

tcggtatgtg aattctatct attttttttc tgatgtgtgc atggatgact catgatcata 10980

ttcttaggta atactgtcaa gcatcaatat ggcaagggcg tttacaagat tgcttcttgg 11040

tctcatatta ctaatgctca cacagttcct ggagaaggta ttatcaaggg acttgccgaa 11100

gtcggcctcc ctcttggtcg tggcttgctt ttgctagcag aaatgtcatc tcaaggtgca 11160

ttaactaagg gtatttacac tgccgaatct gtcaatatgg ctgcgcgcaa caaagatttc 11220

gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga ttttgttgtc 11280

atgtcgctg aagcttggcg taatcatggt catagctggt tcctgtgtga aattgttata 11340

cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tgggggtgcct 11400

aatgagttag ctaactcaca ttaattgctg tgcgctcact gcccgccttc cagtcgggaa 11460

acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagagggc ggtttgctga 11520

ttggggccaaa gacaaaaggg cgacattcaa ccgattgagg gaggggaaggc aaatattgac 11580

ggaaattatt cattaaaggc gaattatcac cgtcaccgac ttgagccatt tgggaattag 11640

agccagcaaa atcaccagta gcaccattac cattagcaag gccggaaacg tcaccaatga 11700

aaccatcgat agcagcaccg taatcagtag cgacagaatc aagtttgcct ttagcgtcag 11760

actgtagcgc gttttcatcg gcattttcgg tcatagcccc cttattagcg tttgccatct 11820

tttcataatc aaaatcaccg gaaccagagc caccaccgga accgcctccc tcagagccgc 11880

caccctcaga accgccaccc tcagagccac caccctcaga gccgccacca gaaccaccac 11940

cagagccgcc gccagcattg acaggaggcc cgatctagta acatagatga caccgcgcgc 12000

gataatttat cctagtttgc gcgctatatt ttgttttcta tcgcgtatta aatgtataat 12060

tgccgggactc taatcataaa aacccatctc ataaataacg tcatgcatta catgttaatt 12120

attacatgct taacgtaatt caacagaaat tatatgataa tcatcgcaag accggcaaca 12180

ggattcaatc ttaagaaact ttattgcaa atgtttgaac gatcggggat catccgggtc 12240

tgtggcgagg actccacgaa aatatccgaa cgagcaaga tatcgcggtg catctcggtc 12300

ttgcctgggc agtcgcccgc gacgccgttg atgtggacgc cgggcccgat catattgtcg 12360

ctcaggatcg tggcgttgtg ctgtgcggcc gttgctgtcg taatgatatc ggcaccttcg 12420
accgcctgtt ccgcagagat cccgtgggcg aagaactcca gcatgagatc cccgcgctgg 12480
aggatcatcc agccggcgtc ccggaaaacg attccgaagc ccaacctttc atagaaggcg 12540
gcggtggaat cgaaatctcg tgatggcagg ttgggcgtcg cttggtcggt catttcgaac 12600
cccagagtcc cgctcagaag aactcgtcaa gaaggcgata gaaggcgatg cgctgcgaat 12660
cgggagcggc gataccgtaa agcacgagga agcggtcagc ccattcgccg ccaagctctt 12720
cagcaatatc acgggtagcc aacgctatgt cctgatagcg gtccgccaca cccagccggc 12780
cacagtcgat gaatccagaa aagcggccat tttccaccat gatattcggc aagcaggcat 12840
cgccatgggt cacgacgaga tcatcgccgt cgggcatgcg cgccttgagc ctggcgaaca 12900
gttcggctgg cgcgagcccc tgatgctctt cgtccagatc atcctgatcg acaagaccgg 12960
cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggtcg aatgggcagg 13020
tagccggatc aagcgtatgc agccgcccga ttgcatcagc catgatggat actttctcgg 13080
caggagcaag gtgagatgac aggagatcct gccccggcac ttcgcccaat agcagccagt 13140
cccttcccgc ttcagtgaca acgtcgagca cagctgcgca aggaacgccc gtcgtggcca 13200
gccacgatag ccgcgctgcc tcgtcctgca gttcattcag ggcaccggac aggtcggctt 13260
tgacaaaaag aaccgggcgc ccctgcgctg acagccggaa cacggcggca tcagagcagc 13320
cgattgtctg ttgtgcccag tcatagccga atagcctctc cacccaagcg gccggagaac 13380
ctgcgtgcaa tccatcttgt tcaatcatgc gaaacgatcc agatccggtg cagattattt 13440
ggattgagag tgaatatgag actctaattg gataccgagg ggaatttatg gaacgtcagt 13500
ggagcatttt tgacaagaaa tatttgctag ctgatagtga ccttaggcga cttttgaacg 13560
cgcaataatg gtttctgacg tatgtgctta gtcattaaa ctccagaaac ccgcggtga 13620

gtggctcctt caacgttgcg gttctgtcag ttccaaacgt aaaacggctt gtcccgcgtc 13680
atcggcgggg gtcataacgt gactccctta attctccgct catgatcaga ttgtcgtttc 13740
ccgccttcag tttaaactat cagtgtttga caggatatat tggcgggtaa acctaagaga 13800
aaagagcgtt tattagaata atcggatatt taaaagggcg tgaaaagggt tatccgttcg 13860
tccatttgta tgtgcatgcc aaccacaggg ttccccagat ctggcgccgg ccagcgagac 13920
gagcaagatt ggccgcgcgc cgaaacgatc cgacagcgcg cccagcacag gtgcgcaggc 13980
aaattgcacc aacgcataca gcgccagcag aatgccatag tgggcgggtga cgtcgttcga 14040
gtgaaccaga tcgcgcagga ggcccggcag caccggcata atcaggccga tgccgacagc 14100
gtcgagcgcg acagtgtca gaattacgat caggggtatg ttgggtttca cgtctggcct 14160
ccggaccagc ctccgctggt ccgattgaac gcgcggattc tttatcactg ataagttggt 14220
ggacatatta tgtttatcag tgataaagtg tcaagcatga caaagttgca gccgaatata 14280
gtgatccgtg ccgccctgga cctgttgaac gaggtcggcg tagacggtct gacgacacgc 14340
aaactggcgg aacggttggg ggttcagcag ccggcgcttt actggcactt caggaacaag 14400
cgggcgctgc tcgacgcact ggccgaagcc atgctggcgg agaatcatac gcattcgggtg 14460
ccgagagccg acgacgactg gcgctcattt ctgatcggga atgcccgag cttcaggcag 14520
gcgctgctcg cctaccgcga tggcgcgcg atccatgccg gcacgcgacc gggcgccaccg 14580
cagatgaaa cgcccgacgc gcagcttcgc ttcctctgcg aggcgggttt ttcggccggg 14640
gacgccgtca atgcgctgat gacaatcagc tacttcactg ttggggccgt gcttgaggag 14700
caggccggcg acagcgatgc cggcgagcgc ggccggcaccg ttgaacaggc tccgctctcg 14760
ccgctgttgc gggccgcgat agacgccttc gacgaagccg gtccggacgc agcgttcgag 14820
cagggactcg cggtgattgt cgatggattg gcgaaaagga ggctcgttgt caggaacgtt 14880

gaaggaccga gaaaggggtga cgattgatca ggaccgctgc cggagcgcaa cccactcact 14940
acagcagagc catgtagaca acatcccctc cccctttcca ccgcgtcaga cgcccgtagc 15000
agcccgtac gggctttttc atgccctgcc ctagcgtcca agcctcacgg ccgcgctcgg 15060
cctctctggc ggccttctgg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg 15120
tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 15180
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 15240
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca 15300
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 15360
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 15420
tgtccgcctt tctcccttcg ggaagcgtgg cgctttttccg ctgcataacc ctgcttcggg 15480
gtcattatag cgattttttc ggtatatcca tccttttttcg cacgatatac aggattttgc 15540
caaaggggtc gtgtagactt tccttggtgt atccaacggc gtcagccggg caggataggt 15600
gaagtaggcc caccgcgag cgggtgttcc ttcttctactg tcccttattc gcacctggcg 15660
gtgctcaacg ggaatcctgc tctgcgaggc tggccggcta ccgccggcgt aacagatgag 15720
ggcaagcggg tggctgatga aaccaagcca accaggaagg gcagcccacc tatcaagggtg 15780
tactgccttc cagacgaacg aagagcgatt gaggaaaagg cggcggcggc cggcatgagc 15840
ctgtcggcct acctgctggc cgtcggccag ggctacaaaa tcacgggcgt cgtggactat 15900
gagcacgtcc gcgagctggc ccgcatcaat ggcgacctgg gccgcctggg cggcctgctg 15960
aaactctggc tcaccgacga cccgcgcacg gcgcggttcg gtgatgccac gatcctcgcc 16020
ctgctggcga agatcgaaga gaagcaggac gagcttggca aggtcatgat gggcgtggtc 16080
cgcccgaggg cagagccatg acttttttag ccgctaaaac ggccgggggg tgccgctgat 16140

tgccaagcac gtcccatgc gctccatcaa gaagagcgac ttcgcggagc tggatgaagta 16200

catcaccgac gagcaaggca agaccgagcg cctttgcgac gctca 16245

<210> 37

<211> 17877

<212> DNA

<213> Artificial

<220>

<223> Promotor

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 37

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttccacacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggg ttccgcccgt 480

ttttcggcca cgcctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctccccggc cgctaacgcg ggccctcccat cccccaggg gctgcgcccc 660

tcggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgcca cggccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactcaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgetgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttcggag actgtcatcac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcggcctg agtttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcacccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcggt gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatcac ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagtgc gtcttggtat 2220

aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc tcaaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc ttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatatTTT tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360
gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgcctcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
ccctgttcac cagcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggctcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggccccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcgc tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatgtcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcg cctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata ttctgctatt 5100

ctggagcttg ttgtttatct cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtctctgggg ctattttcgg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgccctggcaa ctggcggccg gaggacttct gctcgttcca gtagcttttag tgtttgatcc 5460

gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcggggttta acctacttcc tttgggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggcttttctca gccccagatc tggggtcgat cagccgggga tgcatacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgccgtcccc gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatgggtgg tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttt gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggatccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtgggtg caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc	7920
tgtccagatc atgggttgacc ggtgcctgga tcttcctata gaatcatcct tattcggtga	7980
cctagctgat tctggagtga cccagaggggt catgacttga gcctaaaadc cgccgcctcc	8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc	8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag	8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc	8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaacttatt gcatcatcgg	8280
agaatatgga gcttcacga atcaccggca gtaagcgaag gagaatgtga agccaggggt	8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg	8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc	8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc	8520
ctgctttgcc cgggtgtatga aaccggaaaag gccgctcagg agctggccag cggcgcagac	8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc	8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca	8700
aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct	8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt	8820
tatttcccct aagtaagtac ttggtacat ccatactcca tccttcccat cccttattcc	8880
tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctaccccg	8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga	9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtcttt	9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt	9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcggggttcg gcccatctcg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcagc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgac agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaac agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tacttttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttcttttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ttttcgagtt 10800
tttttttttt ttctttgtga aggatttatt gttattggta tccatttttt attggaagac 10860
aagataagtt aatattgatt ttgcttaaag attaaaagga aatcagaaaa cgacaataaa 10920
aaatgtaacg gacaaactat ggtgtcgatt ataagtctaa atccttaaaa aatgacaacg 10980
agttgctttc ctctgaaaac aattcttttg tctttgcaag aaaggtttct tttttgtttg 11040
cttgcatcac ttaaacaatca aatcaaatga aaggaataaa gcagatttga gggcgaataa 11100
ggattttctg gtcaacaaga tgtgagtgac acctaaaggaa ctaaagcca ttcatttggt 11160
ttaaacgac atcaaagatt gatgatcaac aggattgaga gagagaaaaa gaactcgtgt 11220
catttatctc tgttgactga aattttatat ttagaaaaaa tgtcaaatct atagctttag 11280
ctatattaca taacatttga aataataata ataaaaaag acacattaga gacacttttc 11340
aaactctaaa taactgtcta taaacacaaa gaaaacaaag acctctataa caacttatta 11400
gatttttctc gtacttttgt ctaaagatga tgtattcttg ttatcccaca cttctttcat 11460
ttgttcttga tgctactaaa tatacaaaat ttcttttttg caagagatat tattccaaaa 11520
attttcaaaa agaaattttt ttcacaatag cagttgatcg tgtaacccaa agaggttctt 11580
tgttattttg cacttccgct ttgcggtgat gcatattcaa agtaatatat ggaataaaca 11640

acgtgttttaa gcatgaaaga aaggaaacaa aggccgcttt gaacaaatgc ataatatattc 11700

agacaaaaat gatctaaagc aagcagtaaa tcaaacaaga aacattgctg attcgcgтта 11760

gaaaacgata aaagtctaаt aagccactaa gtatacttca atgaactttt tgtatgctta 11820

tggtccaatc agaccaataa tttgtgacca ttcttgaggt ggctttgggtg atgcggaaac 11880

agaaaaaaat tttctcacca atcgatttaa aaaacaattt ctgctttgaa ccaaaacttt 11940

ttttttctct ttaatcatta actttatcaa gtatgtacct accctcaaag tcctcactca 12000

agcacaatta tgctaacatt gttccacctt ctctttagaa atgctgtcga agctgcagtc 12060

aatcagcgtc aaggcccgcc gcgttgaact agcccgcgac atcacgcggc ccaaagtctg 12120

cctgcatgct cagcgggtgct cgttagttcg gctgcgagtg gcagcaccac agacagagga 12180

ggcgctggga accgtgcagg ctgccggcgc gggcgatgag cacagcgccg atgtagcact 12240

ccagcagctt gaccgggcta tcgcagagcg tcgtgcccgg cgcaaacggg agcagctgtc 12300

ataccaggct gccgccattg cagcatcaat tggcgtgtca ggcatgcca tcttcgccac 12360

ctacctgaga tttgccatgc acatgaccgt gggcggcgca gtgccatggg gtgaagtggc 12420

tggcactctc ctcttgggtg ttggtggcgc gctcggcatg gagatgtatg cccgctatgc 12480

acacaaagcc atctggcatg agtcgcctct gggctggctg ctgcacaaga gccaccacac 12540

acctcgact ggacccttg aagccaacga cttgtttgca atcatcaatg gactgcccgc 12600

catgctcctg tgtaccttg gcttctggct gcccaacgtc ctggggggcg cctgctttgg 12660

agcggggctg ggcatacgc tatacggcat ggcataatg tttgtacacg atggcctggt 12720

gcacaggcgc tttcccaccg ggcccatcgc tggcctgcc tacatgaagc gcctgacagt 12780

ggcccaccag ctacaccaca gcggcaagta cgggtggcgc cctgggggta tgttcttggg 12840

tccacaggag ctgcagcaca ttccaggtgc ggcggaggag gtggagcgac tggtcctgga 12900

actggactgg tccaagcggg agaagcttgg cgtaatcatg gtcataagctg tttcctgtgt 12960

gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 13020

cctgggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgtt 13080

tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 13140

gcggtttgcg tattgggcca aagacaaaag ggcgacattc aaccgattga gggaggggaag 13200

gtaaatattg acggaaatta ttcattaaag gtgaattatc accgtcaccg acttgagcca 13260

tttgggaatt agagccagca aaatcaccag tagcaccatt accattagca aggccggaaa 13320

cgtcaccaat gaaaccatcg atagcagcac cgtaatcagt agcgacagaa tcaagtttgc 13380

cttttagcgtc agactgtagc gcgttttcat cggcattttc ggtcatagcc cccttattag 13440

cgtttgccat cttttcataa tcaaaatcac cggaaccaga gccaccaccg gaaccgcctc 13500

cctcagagcc gccaccctca gaaccgccac cctcagagcc accaccctca gagccgccac 13560

cagaaccacc accagagccg ccgccagcat tgacaggagg cccgatctag taacatagat 13620

gacaccgcgc gcgataattt atcctagttt gcgcgctata ttttgttttc tatcgcgat 13680

taaatgtata attgcgggac tctaatcata aaaaccatc tcataaataa cgtcatgcat 13740

tacatgttaa ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca 13800

agaccggcaa caggattcaa tcttaagaaa ctttattgcc aaatgtttga acgatcgggg 13860

atcatccggg tctgtggcgg gaactccacg aaaatatccg aacgcagcaa gatatcgcg 13920

tgcatctcgg tcttgccctg gcagtcgccg ccgacgccgt tgatgtggac gccgggcccc 13980

atcatattgt cgctcaggat cgtggcggtg tgcttgctcg ccgttgctgt cgtaatgata 14040

tcggcacctt cgaccgcctg ttccgcagag atcccgtggg cgaagaactc cagcatgaga 14100

tccccgcgct ggaggatcat ccagccggcg tcccggaaaa cgattccgaa gcccaacctt 14160
tcatagaagg cggcgggtgga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg 14220
gtcatttcga accccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga 14280
tgcgctgcga atcgggagcg gcgataccgt aaagcacgag gaagcgggtca gccattcg 14340
cgccaagctc ttcagcaata tcacgggtag ccaacgctat gtctgatag cggtcgcga 14400
caccagccg gccacagtcg atgaatccag aaaagcggcc atttccacc atgatattcg 14460
gcaagcaggc atcgccatgg gtcacgacga gatcatcgcc gtcgggcatg cgcgccctga 14520
gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc ttcgtccaga tcatcctgat 14580
cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat gcgatgttcc gcttggtggt 14640
cgaatgggca ggtagccgga tcaagcgtat gcagccgccg cattgcatca gccatgatgg 14700
atactttctc ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgcccc 14760
atagcagcca gtcccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc 14820
ccgtcggtggc cagccacgat agccgcgctg cctcgctctg cagttcatcc agggcaccgg 14880
acaggtcggt cttgacaaaa agaaccgggc gccctgctgc tgacagccgg aacacggcgg 14940
catcagagca gccgattgtc tgttgtgccc agtcatagcc gaatagcctc tccaccaag 15000
cggccggaga acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg 15060
tgcagattat ttggattgag agtgaatatg agactctaatt tggataccga ggggaattta 15120
tggaacgtca gtggagcatt tttgacaaga aatatttgct agctgatagt gaccttaggc 15180
gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 15240
acccgcggct gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaaacggc 15300
ttgtcccgcg tcatcggcgg gggtcataac gtgactccct taattctccg ctcatgatca 15360

gattgtcggtt tcccgccttc agtttaaact atcagtggtt gacaggatat attggcgggt 15420

aaacctaaga gaaaagagcg tttattagaa taatcggata tttaaaaggg cgtgaaaagg 15480

tttatccgtt cgtccatttg tatgtgcatg ccaaccacag ggttccccag atctggcgcc 15540

ggccagcgag acgagcaaga ttggccgccc cccgaaacga tccgacagcg cgcccagcac 15600

aggtgcgag gcaaattgca ccaacgcata cagcgccagc agaatgcat agtgggagg 15660

gacgtcggtt gagtgaacca gatcgcgag gagggccggc agcaccggca taatcaggcc 15720

gatgccgaca gcgtcgagcg cgacagtgtt cagaattacg atcaggggta tggtgggtt 15780

cacgtctggc ctccggacca gcctccgtg gtccgattga acgcgcggtt tctttatcac 15840

tgataagttg gtggacatat tatgtttatc agtgataaag tgtcaagcat gacaaagttg 15900

cagccgaata cagtgatccg tgccgccctg gacctgttga acgaggtcgg cgtagacggg 15960

ctgacgacac gcaaactggc ggaacgggtt ggggttcagc agccggcgct ttactggcac 16020

ttcaggaaca agcggggcgct gctcgacgca ctggccgaag ccatgctggc ggagaatcat 16080

acgcattcgg tgccgagagc cgacgacgac tggcgctcat ttctgatcgg gaatgcccgc 16140

agcttcaggc aggcgctgct cgcctaccgc gatggcgcg ccatccatgc cggcacgcga 16200

ccgggcgcac cgcagatgga aacggccgac gcgcagctt gcttcctctg cgaggcgggt 16260

ttttcggccg gggacgccgt caatgcgctg atgacaatca gctacttcac tggtggggcc 16320

gtgcttgagg agcaggccgg cgacagcgat gccggcgagc gcggcgccac cgttgaacag 16380

gctccgtct cgcgctgtt gcgggccgcg atagacgcct tcgacgaagc cgggtccggac 16440

gcagcgttcg agcagggact cgcggtgatt gtcgatggat tggcgaaaag gaggtcgtt 16500

gtcaggaacg ttgaaggacc gagaaagggt gacgattgat caggaccgct gccggagcgc 16560

aaccactca ctacagcaga gccatgtaga caacatcccc tcccccttc caccgcgtca 16620

gacgcccgtg gcagcccgtt acgggctttt tcatgccctg ccctagcgtc caagcctcac 16680

ggccgcgctc ggcctctctg gcggccttct ggcgctcttc cgcttcctcg ctactgact 16740

cgctgcgctc ggtcgttcgg ctgcggcgag cggatcagc tcaactcaaag gcggaatac 16800

ggttatccac agaatacagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 16860

aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg 16920

acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 16980

gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttcgg accctgccgc 17040

ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgcttttc cgctgcataa 17100

ccctgcttcg gggtcattat agcgattttt tcggtatata catccttttt cgcacgatat 17160

acaggatttt gccaaagggc tcgtgtagac tttccttggt gtatccaacg gcgtcagccg 17220

ggcaggatag gtgaagtagg cccacccgcg agcgggtggt ccttcttcac tgtcccttat 17280

tcgcacctgg cgggtgctca cgggaatcct gctctgcgag gctggccggc taccgccggc 17340

gtaacagatg agggcaagcg gatggctgat gaaaccaagc caaccaggaa gggcagccca 17400

cctatcaagg tgtactgcct tccagacgaa cgaagagcga ttgaggaaaa ggcggcggcg 17460

gccggcatga gcctgtcggc ctacctgctg gccgtcggcc agggctacaa aatcacgggc 17520

gtcgtggact atgagcacgt ccgcgagctg gccgcacat atggcgacct gggccgcctg 17580

ggcggcctgc tgaaactctg gtcaccgac gaccgcgca cggcgcggtt cggatgatgc 17640

acgatcctcg ccctgctggc gaagatcgaa gagaagcagg acgagcttgg caaggtcatg 17700

atgggcgtgg tccgcccag ggcagagcca tgactttttt agccgctaaa acggccgggg 17760

ggtgcgcgtg attgccaagc acgtcccat gcgtccatc aagaagagcg acttcgcgga 17820

gctggtgaag tacatcacgc acgagcaagg caagaccgag cgcctttgcg acgctca 17877

<210> 38
<211> 17238
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 38
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtgga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttccccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtgta cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcggcca cgcctaacct gtcttttaac ctgcttttaa accaatatTT ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgT gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctccccggc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc 660

tcggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgC cattgcccgg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgcgA cccccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggTgccgg gcagtgaggg cggcggcctg 840

ggTggcgggc tgccttcac ttcggccgTc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggTcgcgg gtgccgtgct cgtgttcggg 960

ggTgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgctt aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgcgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggTgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcggcctg agtttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220

aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtctctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcttttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc tcaaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttggggagaag cggcagggcg gacaagtgg atgacattgc 2940

cttctgctc cggctgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggtg agtcggggca atcccgcaag 3360
gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggtcc agcaagctac ggccaagatc gagcgcgaca 3540
gggtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaana accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgtctctg 3840
ccctgttcac cagcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggt gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccggtgc caggctctga tcgacgagga aatcgctcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata tttegetatt 5100

ctggagcttg ttgtttatct cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtccctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgaggctcg gcctgatcgg 5520

agcggggttta acctacttcc tttgggtccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gccccagatc tggggtcgat cagccgggga tgcacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcgggttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgctga cgccgtcccg gactgatggg ctgcctgtat 6060

cgagtgggtga ttttggtccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggt 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tggttaacgc caggggttttc ccagtcacga cggttgtaaaa cgacggccag tgaattcgag	6780
ctcgggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg	6840
cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca	6900
gctcattaag agccagttca tgggcgttgg catgatggcc gtcattgcac tgtacttcaa	6960
gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt	7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc	7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag	7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag	7200
cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg	7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg	7320
tcctttgggt ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt	7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt	7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg	7500
gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc	7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga	7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct	7680
gtaaggatth cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc	7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt	7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa	7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatt	7920
tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcggtga	7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaatc cgccgcctcc	8040
accattttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc	8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag	8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc	8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg	8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt	8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg	8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc	8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc	8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac	8580
cgggaaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc	8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca	8700
aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct	8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt	8820
tatttcccct aagtaagtac tttgctacat ccatactcca tccttcccat cccttattcc	8880
tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccgc	8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga	9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctcgtgcttt	9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt	9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt	9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg	9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgagggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcttgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcy acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc aactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcggttg tgcgatgct agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctctgttact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ctaccgcttg 10800
gaccagtcca gttccaggac cagtcgctcc acctcctccg ccgcacctgg aatgtgctgc 10860
agtcctgtg gacccaagaa cataccccag ggcgcgccac cgtacttgcc gctgtggtgt 10920
agctggtggg ccactgtcag gcgcttcatg tagggcaggc cagcgatggg cccggtggga 10980
aagcgctgt gcaccaggcc atcgtgtaca aacatatatg ccatgccgta tagcgtgatg 11040
cccagccccg ctccaaagca ggccgcccc aggacgttg gcagccagaa gccaaaggta 11100
cacaggagca tggcgggcag tccattgatg attgcaaaca agtcgttggc ttcaaagggt 11160
ccagtgcgag gtgtgtggtg gctcttgtgc agcagccagc ccagaggcga ctcatgccag 11220
atggctttgt gtgcatagcg ggcatacatc tccatgccga gcgcgccacc aaccaccaag 11280
aggagagtgc cagccacttc acccctatggc actgcgccgc ccacggtcac gtgcatggca 11340
aatctcaggt aggtggcgaa gatggcaatg cctgacacgc caattgatgc tgcaatggcg 11400
gcagcctggt atgacagctg ctcccgtttg cgccgggcac gacgctctgc gatagcccgg 11460
tcaagctgct ggagtgtac atcggcgctg tgctcatcgc ccgcgccggc agcctgcacg 11520
gttcccagcg cctcctctgt ctgtggtgct gccactcgca gccgaactaa cgagcaccgc 11580
tgagcatgca ggcagacttt gggccgcgtg atgtcgcggt ctagttcaac gcggcgggcc 11640
ttgacgctga ttgactgcag cttcgacagc atagagataa aataaaaaga gaagaaaaga 11700
aagtttgtac aatttctttt tgtttatata acatacacgc tatgtcaaca tttagaataa 11760

gggggaaaaa atcttccatc atattcgaat gcacaagatt atttctttgt tcgctctttt 11820
tggtcgggtc atcgagattt agagtgtaat caaagatact gtcattctga gagcggtgca 11880
caggctgctg tttgccaaat tggatgtttg ccgaattagt aaaatacgca agcatttctt 11940
acctttccgc tcccttttcc taattctccc aaagactaaa tgaggaaaga taaaggacaa 12000
agaaaatgta aagacaaaga aattgaaaac gatataaact tgcagcacgt aagaccaaag 12060
caaattggta actattcttg tgtacaaaca tgtataaaaa aaaacttttt tttgctcctg 12120
gaggacaaaa tttcaaactc cttgaagaag attgcttgta tatctatcat atgcatatat 12180
catatcgatg gaaaaagaaa gtcaggcatg tatttataaa aagaagaatg tgccatgctt 12240
ccgaatttct tttcactttc ttttccttat ctattttaat ctcaagcttg gcgtaatcat 12300
ggtcatagct gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacgag 12360
ccggaagcat aaagtgtaaa gcctgggggtg cctaatgagt gagctaactc acattaattg 12420
cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 12480
tcggccaacg cgcggggaga ggcggtttgc gtattgggcc aaagacaaaa gggcgacatt 12540
caaccgattg agggagggaa ggtaaatatt gacggaaatt attcattaaa ggtgaattat 12600
caccgtcacc gacttgagcc atttgggaaat tagagccagc aaaatcacca gtagcaccat 12660
taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12720
tagcgacaga atcaagtttg ctttagcgt cagactgtag cgcgttttca tcggcatttt 12780
cggtcatagc ccccttatta gcgtttgcca tcttttcata atcaaaatca ccggaaccag 12840
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgcca ccctcagagc 12900
caccaccctc agagccgcca ccagaaccac caccagagcc gccgccagca ttgacaggag 12960
gcccgatcta gtaacataga tgacaccgcg cgcgataatt taccctagtt tgccgcgctat 13020

atTTTgtttt ctatcgcgta ttaaAtgtat aattgcggga ctctaAtcat aaaaacccat 13080
ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13140
aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13200
caaAtgtttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatattcc 13260
gaacgcagca agatatcgcg gtgcatctcg gtcttgccctg ggcagtcgcc gccgacgccg 13320
ttgatgtgga cgccggggccc gatcatattg tcgctcagga tcgtggcggtt gtgcttgctg 13380
gccgttgctg tcgtaatgat atcggcacct tcgaccgcct gttccgcaga gatcccgtgg 13440
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccggaaa 13500
acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13560
aggttggggcg tcgcttggtc ggtcatttcg aaccccagag tcccgtctag aagaactcgt 13620
caagaaggcg atagaaggcg atgcgctgcg aatcgggagc ggcgataccg taaagcacga 13680
ggaagcggtc agccattcg ccgccaagct cttcagcaat atcacgggta gccaacgcta 13740
tgtcctgata gcggctccgc acaccagcc ggccacagtc gatgaatcca gaaaagcggc 13800
cattttccac catgatattc ggcaagcagg catcgccatg ggtcacgacg agatcatcgc 13860
cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13920
cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtacgt gctcgctcga 13980
tgcgatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 14040
gcattgcac agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 14100
cctgccccgg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga 14160
gcacagctgc gcaaggaacg cccgtcgtgg ccagccacga tagccgcgct gcctcgtcct 14220
gcagttcatt cagggcaccg gacaggtcgg tcttgacaaa aagaaccggg cgcccctgcg 14280

ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14340
cgaatagcct ctccacccaa gcggccggag aacctgcgtg caatccatct tgttcaatca 14400
tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460
ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatttgc 14520
tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14580
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14640
cagttccaaa cgtaaaacgg cttgtccgc gtcacggcg ggggtcataa cgtgactccc 14700
ttaattctcc gctcatgac agattgtcgt ttccgcctt cagtttaaac tatcagtgtt 14760
tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataacggat 14820
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaccaca 14880
gggttcccca gatctggcg cgccagcga gacgagcaag attggccgcc gcccgaaacg 14940
atccgacagc gcgccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag 15000
cagaatgcca tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg 15060
cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15120
gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15180
aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtataaa 15240
gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15300
aacgaggtcg gcgtagacgg tctgacgaca cgcaaactgg cggaacgggtt gggggttcag 15360
cagccggcg tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15420
gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15480
tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15540

cgcatccatg ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt 15600
cgcttcctct gcgaggcggg tttttcggcc ggggacgccc tcaatgcgct gatgacaatc 15660
agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15720
cgcgggcgga ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc 15780
ttcgacgaag ccgggtccga cgcagcgctt gagcaggggac tcgcggtgat tgatcgatgga 15840
ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15900
tcaggaccgc tgccggagcg caaccactc actacagcag agccatgtag acaacatccc 15960
ctcccccttt ccaccgctc agacgcccggt agcagccccgc tacgggcttt ttcattgcct 16020
gccctagcgt ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt 16080
ccgcttcctc gctcactgac tcgctgcgct cggctcgctc gctgcggcga gcggtatcag 16140
ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16200
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgctg ctggcgcttt 16260
tccataggct ccgccccctt gacgagcatc aaaaaatcg acgctcaagt cagagggtggc 16320
gaaacccgac aggactataa agataccagg cgtttcccc tggaaagctcc ctgctgcgct 16380
ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 16440
tggcgctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat 16500
ccatcccttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16560
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt 16620
tccttcttca ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga 16680
ggctggccgg ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16740
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16800

attgaggaaa aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16860
cagggctaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcac 16920
aatggcgacc tgggcccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc 16980
acggcgcggt tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag 17040
gacgagcttg gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt 17100
tagccgctaa aacggccggg ggggtgcgcgt gattgccaag cacgtcccca tgcgctccat 17160
caagaagagc gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17220
gcgcctttgc gacgctca 17238

<210> 39

<211> 17238

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 39

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtgga tacctcgcg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttccccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt	480
ttttcgcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaa cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcgcccc	660
tcggccgga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcatcagc ccgagcgga cgcccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcgccctg	840
ggtggcggcc tgcccttcac ttcgccgctc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140
agataatata tcttttatat agaagatct gccgtatgta aggatttcag ggggcaaggc	1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga	1260

ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaaggata taagctggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggtcctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggc attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttgcgga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgtctcg 3840
ccctgttcac cagcgcaac aagaaaatcc cgcgcgaggc gctgcaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgccg tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgetgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgcg 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccg gctcaagctg gaaaccttc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctg 4500
tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcg tctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgcg gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgtgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccga cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata ttctgctatt 5100

ctggagcttg ttgtttatctt cgggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatagc 5220

attgatggcg gtcctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgcttgga caa ctggcggccg gaggacttct gctcggtcca gtagctttag tggttgatcc 5460

gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc ttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcaact tctaaagaaa tagcgccact cagcttctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgccgtcccc gactgatggg ctgcctgtat 6060

cgagtgggta ttttggtccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatcc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440

tcgaaataca tccgatgtc gaaggcttgg ggcacctgcg ttgggtgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaata 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga ccagaggggt catgacttga gcctaaaata cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgcggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640

tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggctcgttg gtcagtccaa catttgttg catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac tttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggccaagaat ctcgtgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacgggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggg 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag gtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct cgggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatataattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800
taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860
tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920
ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca aagatactgt 10980
catctcgaga gcgttgca ca ggctgctgtt tgccaaattg gatgtttgcc gaattagtaa 11040
aatacgcaag catttcttac ctttccgctc ctttttcta attctccaa agactaaatg 11100
aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga tataaacttg 11160
cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaacatg tataaaaaaa 11220
aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat tgcttgata 11280
tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta tttataaaaa 11340

gaagaatgtg ccatgcttcc gaatttcttt tcaactttctt ttccttatct attttaatct 11400
catgctgtcg aagctgcagt caatcagcgt caaggcccg cgcggtgaac tagcccgca 11460
catcacgcgg cccaaagtct gcctgcatgc tcagcgggtgc tcgttagttc ggctgcgagt 11520
ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgccggcg cgggcgatga 11580
gcacagcgcc gatgtagcac tccagcagct tgaccgggct atcgagagc gtcgtgcccc 11640
gcgcaaacgg gagcagctgt cataccaggc tgccgccatt gcagcatcaa ttggcgtgtc 11700
aggcattgcc atcttcgcca cctacctgag atttgccatg cacatgaccg tgggcggcg 11760
agtgccatgg ggtgaagtgg ctggcactct cctcttgggtg gttgggtggcg cgctcggcat 11820
ggagatgtat gcccgtatg cacacaaagc catctggcat gagtcgcctc tgggctggct 11880
gctgcacaag agccaccaca cacctcgac tggaccctt gaagccaacg acttgtttgc 11940
aatcatcaat ggactgccc ccatgctcct gtgtacctt ggcttctggc tgcccaacgt 12000
cctggggggcg gcctgctttg gagcgggggt gggcatcac ctatacggca tggcatatat 12060
gtttgtacac gatggcctgg tgcacaggcg ctttcccacc gggcccatcg ctggcctgcc 12120
ctacatgaag cgctgacag tggccacca gctacaccac agcggcaagt acggtggcg 12180
gccctgggggt atgttcttgg gtccacagga gctgcagcac attccagggt cggcggagga 12240
gggtggagcga ctggctcctgg aactggactg gtccaagcgg tagaagcttg gcgtaatcat 12300
ggcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag 12360
ccggaagcat aaagtgtaaa gcctgggggtg cctaagtgt gagctaactc acattaattg 12420
cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 12480
tcggccaacg cgcggggaga ggcggtttgc gtattgggcc aaagacaaaa gggcgacatt 12540
caaccgattg agggagggaa ggtaaatatt gacggaaatt attcattaaa ggtgaattat 12600

caccgtcacc gacttgagcc atttggaat tagagccagc aaaatcacca gtagcaccat 12660
taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12720
tagcgacaga atcaagtttg cctttagcgt cagactgtag cgcgttttca tcggcatttt 12780
cggatcatagc ccccttatta gcgtttgcca tcttttcata atcaaaatca ccggaaccag 12840
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgcca ccctcagagc 12900
caccaccctc agagccgcca ccagaaccac caccagagcc gccgccagca ttgacaggag 12960
gcccgatcta gtaacataga tgacaccgag cgcgataatt taccctagtt tgccgctat 13020
attttgtttt ctatcgcgta ttaaatgtat aattgcggga ctctaatacat aaaaacccat 13080
ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13140
aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13200
caaagtgttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13260
gaacgcagca agatatcgcg gtgcatctcg gtcttgccctg ggcagtcgcc gccgacgccg 13320
ttgatgtgga cgccggggccc gatcatattg tcgctcagga tcgtggcggt gtgcttgctg 13380
gccgttgctg tcgtaatgat atcggcacct tcgaccgcct gttccgcaga gatcccgtag 13440
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccgga 13500
acgattccga agcccaacct ttcatagaag gcggcggttg aatcgaaatc tcgtgatggc 13560
agggtgggag tcgcttggtc ggtcatttcg aaccccagag tcccgctcag aagaactcgt 13620
caagaaggcg atagaaggcg atgcgctgag aatcgggagc ggcgataccg taaagcacga 13680
ggaagcggtc agccattcg ccgccaagct cttcagcaat atcacgggta gccaacgcta 13740
tgtcctgata gcgggtccgc acaccagcc ggccacagtc gatgaatcca gaaaagcggc 13800
cattttccac catgatattc ggcaagcagg catcgccatg ggtcacgacg agatcatcgc 13860

cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc cctgatgct 13920
cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtacgt gctcgctcga 13980
tgcgatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 14040
gcattgcac agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 14100
cctgccccgg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga 14160
gcacagctgc gcaaggaacg cccgtcgtgg ccagccacga tagccgcgct gcctcgctct 14220
gcagttcatt cagggcacccg gacaggtcgg tottgacaaa aagaaccggg cgccctgcg 14280
ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14340
cgaatagcct ctccacccaa ggggccggag aacctgcgtg caatccatct tgttcaatca 14400
tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460
ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatgtgc 14520
tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14580
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14640
cagttccaaa cgtaaaacgg ctgtgccgc gtcacggcg ggggtcataa cgtgactccc 14700
ttaattctcc gtcctgatc agattgtcgt ttccgcctt cagtttaaac tatcagtgtt 14760
tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataatcggat 14820
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gcccaaccaca 14880
gggttcccca gatctggcg cggccagcga gacgagcaag attggccgcc gcccgaaacg 14940
atccgacagc gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag 15000
cagaatgcca tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg 15060
cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15120

gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15180
aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtataaa 15240
gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15300
aacgaggtcg gcgtagacgg tctgacgaca cgcaaactgg cggaacgggtt gggggttcag 15360
cagccggcgc tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15420
gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15480
tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15540
cgcattcatg ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgagctt 15600
cgcttctct gcgagggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc 15660
agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15720
cgccggcgga ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc 15780
ttcgacgaag ccggtccgga cgcagcgctc gagcaggac tcgcggtgat tgtcgatgga 15840
ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15900
tcaggaccgc tgccggagcg caaccactc actacagcag agccatgtag acaacatccc 15960
ctccccctt ccaccgctc agacgcccg agcagccgc tacgggcttt ttcatgccct 16020
gccctagcgt ccaagcctca cggccgcgct cggcctctct ggccgccttc tggcgctctt 16080
ccgcttctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag 16140
ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16200
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgctt ctggcgcttt 16260
tccatagget ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagaggtggc 16320
gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc ctctgcgct 16380

ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 16440
tggegcctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat 16500
ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16560
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gcccaccgc gagcgggtgt 16620
tcctttcttca ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga 16680
ggctggccgg ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16740
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16800
attgaggaaa aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16860
cagggttaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcac 16920
aatggcgacc tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgaccgcgc 16980
acggcgcggt tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag 17040
gacgagcttg gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt 17100
tagccgctaa aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat 17160
caagaagagc gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17220
gcgcctttgc gacgctca 17238

<210> 40

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>

<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>

<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 40

gatctttcga cactgaaata cgtcgagcct gtcctcgcttg gaagcggcga ggagcctcgt	60
cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggccgctc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct	240
tcgggaagcc agcgactggg gacctccagc gtccctttaa ggctgccaac agctttctca	300
gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtgggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggc cgttgcacia tgcaagtcaa acgttgctag cagttccagg	780
tggaaatgta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840

caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gattaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgaccc agagggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccgggtga ctctttctgg catgaggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcctcggaaga atatggagct 1500

tcctcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgatc tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgctc caaatatcgt gcctctcctg ctttgcccg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctgt ccgcccgggtg tctcggcggg gttgacaagg tcgttgcgctc 1920

agtcacaacat ttgttgccat attttctctg tctccccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat tccccctaag 2040

taagtacttt gctacatcca tactccatcc tcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctccggagggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcgctg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340
tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgtc tgacattggg 2400
gaattcagcg agagcctgac ctattgcac tcccgccgtg cacaggggtgt cacgttgcaa 2460
gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgaggaggc catggatgcg 2520
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580
ggccaataca ctacatggcg tgatttcata tgccgcgattg ctgatcccca tgtgtatcac 2640
tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg cccgaagtc cggcacctcg tgcacgcgga ttccggctcc 2760
aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820
ttcgggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgagg atcgccgcgg 2940
ctccggggcgt atatgtccg cattgggtctt gaccaactct atcagagctt ggttgacggc 3000
aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtccg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgcccg cgaagcgcg ccgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctcgcccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagt cttctagtg 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600
atatttgtct ctaattgtac tttgacatgc tctctttctt tactctgata gcttgactat 3660
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780
gatggtatac aatagtaacc atgcatgggt gcctagtga tgctccgtaa caccaatac 3840
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacagggtaca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020
agaaagtga aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080
ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140
ggagtttgaa attttgtcct ccaggagcaa aaaaagttt ttttttatac atgtttgtac 4200
acaagaatag ttaccaatth gctttggtct tacgtgctgc aagtttatat cgttttcaat 4260
ttctttgtct ttacattttc tttgtccttt atctttcctc atttagtctt tgggagaatt 4320
aggaaaaggg agcggaaagg taagaaatgc ttgcgtattt tactaattcg gcaaacatcc 4380
aatttggaac acagcagcct gtgcaacgct ctcgagatga cagtatcttt gattacactc 4440
taaattctga tgacccgacc aaaaagagcg acaaagaaa taatcttgtg cattcgaata 4500
tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560
aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtcgaa 4620

gctgcagtca atcagcgtca agggccgccc cgttgaacta gcccgcgaca tcacgcggcc 4680

caaagtctgc ctgcatgctc agcgggtgctc gttagttcgg ctgcgagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgagagcgt cgtgcccggc gcaaacggga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttcgccacc tacctgagat ttgccatgca catgaccgtg ggcggcgag tgccatgggg 4980

tgaagtggct ggcactctcc tcttggtggt tgggtggcgcg ctcggcatgg agatgtatgc 5040

ccgctatgca cacaagcca tctggcatga gtcgcctctg ggctggctgc tgcacaagag 5100

ccaccacaca cctcgactg gacccttga agccaacgac ttgtttgcaa tcatcaatgg 5160

actgcccgc atgtcctgt gtaccttgg cttctggctg cccaacgtcc tgggggcggc 5220

ctgctttgga gcggggctgg gcatcacgct atacggcatg gcataatgt ttgtacacga 5280

tggcctggtg cacaggcgt tccccaccg gcccatcgct ggcctgccct acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctggggtat 5400

gttcttgggt ccacaggagc tgcagcacat tccagggtcg gcggaggagg tggagcgact 5460

ggtcctggaa ctggactggt ccaagcggta gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580

cgcccgtagc ctaggtgatc aaagactggt gggcttgtct ctgaagcttg taggaaaggc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cgggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760

gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaga ttgtctgtt 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880

gattaagagt ttttttggtg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

aaagagacaa cacatacaat ataaaatcat tgtcttttaa tgtaattta gtagagtgt 6000

aagcctgcat ttttttgta cgcataaaca atgaattcac cccgcttctg gtttttaa 6060

aattatgtca aactagggaa aattcttttt tttctcttcg ttcttttttt ggcttggtgt 6120

ggagtcacag gcttgtcttc agattgatag aggttgtata cactcaacag agcaatcttg 6180

gcacgttcgc ttctttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt 6240

acagatocaa ttgcagcata tacttggcaa gcagctcgta caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tctacaagc 6360

ttcagagaca agcccaacag tctttgatca cctaggctac gggcggtgcc ttgctgtatt 6420

tgttctgttt cttcttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480

gagccatact acagcataaa aggatacgtt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac aaaaaaaaaag aaatttaaga tgagtaggac ttccattctc tcaaaaattt 6600

tattcaatcc ataatgaat tatttttggg caaaaaagaa agattatgcc tgattttctc 6660

tatttttttt ttttttacia ctccaccaat actttctagc ccagcttggc gtaatcatgg 6720

tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia catagagcc 6780

ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg 6840

ttgcgctcac tgcccgtttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900

ggccaacgcg cggggagagg cggtttgcgt attgggcaa agacaaaagg gcgacattca 6960

accgattgag ggagggaagg taaatattga cggaaattat tcattaaagg tgaattatca 7020

ccgtcaccga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140

gcgacagaat caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg 7200
gtcatagccc ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag 7260
ccaccaccgg aaccgcctcc ctgagagccg ccaccctcag aaccgccacc ctgagagcca 7320
ccaccctcag agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggg 7380
ccgatctagt aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat 7440
tttgttttct atcgcgattt aaatgtataa ttgcggggact ctaatcataa aaacccatct 7500
cataaataac gtcattgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560
ttatatgata atcatcgcaa gaccggcaac aggattcaat ctttaagaaac tttattgcca 7620
aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680
acgcagcaag atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt 7740
gatgtggacg ccggggcccga tcatattgtc gtcaggatc gtggcggtgt gcttgtcggc 7800
cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccggtggc 7860
gaagaactcc agcatgagat cccgcgctg gaggatcatc cagccggcgt cccggaaaac 7920
gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980
gttgggcgtc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040
agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100
aagcggtcag cccattcgcc gccaaagtct, tcagcaatat cacgggtagc caacgctatg 8160
tcctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220
ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg 8280
tcgggcatgc gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct 8340
tcgtccagat catcctgatc gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg 8400

cgatgtttcg cttggtgggc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460
attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520
tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgc aacgtcgagc 8580
acagctgcgc aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctgcctctgc 8640
agttcattca gggcaccgga caggtcgggc ttgacaaaaa gaaccgggcg cccctgcgct 8700
gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg 8760
aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820
cgaaacgac cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880
ggataccgag gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940
gctgatagtg accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000
agctcattaa actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca 9060
gttccaaacg taaaacggct tgtcccgcgt catcgccggg ggtcataacg tgactccctt 9120
aattctccgc tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 9180
acaggatata ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat 9240
ttaaaagggc gtgaaaaggc ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300
gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgccgc ccgaaacgat 9360
ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420
gaatgccata gtgggcgggtg acgtcgttcg agtgaaccag atcgcgcagg agggccggca 9480
gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga 9540
tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600
cgcgccggatt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa 9720
cgaggtcggc gtagacgggc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780
gccggcgctt tactggcact tcaggaacaa gccggcgctg ctcgacgcac tggccgaagc 9840
catgctggcg gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt 9900
tctgatcggg aatgcccga gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg 9960
catccatgcc ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg 10020
cttcctctgc gaggcgggtt ttccggccgg ggacgccgtc aatgcgctga tgacaatcag 10080
ctacttcact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140
cggcggcacc gttgaacagg ctccgctctc gccgctgttg cgggcgcga tagacgcctt 10200
cgacgaagcc ggtccggaag cagcgttcga gcagggactc gcggtgattg tcgatggatt 10260
ggcgaaaagg aggcctcgtt tcaggaacgt tgaaggaccg agaaaggggtg acgattgate 10320
aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380
ccccctttcc accgcgtcag acgcccgtag cagcccgcta cgggcttttt catgccctgc 10440
cctagcgtcc aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500
gcttcctcgc tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 10560
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740
aaccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 10800
cctgttccga ccctgccgct taccggatac ctgtccgctt ttctcccttc gggaagcgtg 10860
gcgcttttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc 10920

atcctttttc gcacgatata caggattttg ccaaaggggtt cgtgtagact ttccttggtg 10980
tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc 11040
cttcttcact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100
ctggccgggt accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160
aaccaggaag ggcagccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11220
tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280
gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340
tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 11400
ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460
cgagcttggc aaggatcatga tgggcgtggt ccgcccaggg gcagagccat gactttttta 11520
gccgctaaaa cgcccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580
agaagagcga cttcgcggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc 11640
gcctttgcga cgctcaccgg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc 11700
cctgcaaacg cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt 11760
tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820
tgagggggccg actcaccgg cgccggcgtt acagatgagg ggcaggctcg atttcggccg 11880
gcgacgtgga gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc 11940
ccacagatga tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc 12000
gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060
tgagggggcg acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc 12120
aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 12180

atatttataa accttgtttt taaccagggc tgcgccctgt gcgctgacc gcgcacgccg 12240
aaggggggtg ccccccttc tcgaaccctc ccggcccgct aacgcgggcc tcccatcccc 12300
ccaggggctg cgccttcg cgcgaacgg cctcaccca aaaatggcag cgctggcagt 12360
ccttgccatt gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc 12420
cggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480
tgagggcggc ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga 12540
cttcattggcg gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600
cgtgctcgtg ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt 12660
ataccgaggt atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat 12720
ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat 12780
attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840
tttcaggggg caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca 12900
taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960
ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc 13020
gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg 13080
tgccagggtg tgctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140
gattacgtgc agctttccct tcaggcgga ttcatacagc ggccagccat ccgtcatcca 13200
tatcaccacg tcaaagggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg 13260
ttcaccgaat acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca 13320
gcgctggcgc gatttagccc cgacatagcc ccactgttcg tccatttccg cgcagacgat 13380
gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga 13440

cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccgga tccaacgcca 13500
ttcatggcca tatcaatgat tttctgggtgc gtaccgggtt gagaagcggg gtaagtgaac 13560
tgcagttgcc atgtttttacg gcagtgagag cagagatagc gctgatgtcc ggcggtgctt 13620
ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680
gccactggag cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc 13740
cataattgtg gtttcaaaat cggctccgctc gatactatgt tatacgccaa ctttgaaaac 13800
aactttgaaa aagctgtttt ctggatttta aggtttttaga atgcaaggaa cagtgaattg 13860
gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920
ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980
accgctgcgt aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040
aaaatgaaaa cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg 14100
tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaaggctc 14160
tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgctc 14220
tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280
agtgcacag gctctttcac tccatcgaca tatcggaattg tccctatacg aatagcttag 14340
acagccgctt agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg 14400
aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga 14460
cggaaaagcc cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520
ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580
agtggatatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt 14640
atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700

atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760
caggagcgca ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc 14820
aagtatttgg gcaaggggtc gctgggtattc gtgcagggca agattcggaa taccaagtac 14880
gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940
gacaccaagg caccaggcgg gtcaaatacag gaataagggc acattgcccc ggctgtgagtc 15000
ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060
gaactgatcg acgcgggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc 15120
atgcgtgcgc cccgcgaaac cttccagtcg gtcggctcga tggccagca agctacggcc 15180
aagatcgagc gcgacagcgt gcaactggct cccctgccc tgcccgcgc atcgccgcgc 15240
gtggagcgtt cgcgtcgtct cgaacaggag gcggcaggtt tggcgaagtc gatgaccatc 15300
gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa 15360
caggtcagcg aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420
atgcagcttt ccttggtcga tattgcgccg tggccggaca cgatgcgagc gatgccaaac 15480
gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540
caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag 15600
ctgcgggccc acgatgacga actggtgtgg cagcaggtgt tggagtacgc gaagcgcacc 15660
cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggctg 15720
atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780
atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgctgctgct gcaccgcttc 15840
cgctctctgg accgtggcaa gaâaacgtcc cgttgccagg tctgatcga cgaggaaatc 15900
gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960

tgcgccgacgg cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgcgtc 16020
aagctggaaa ccttcgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080
gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctgggtgga acacgcctgg 16140
gtcaatgatg acctgggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200
ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260
tgcttcgctc agtatcgctc gggacgcacg gcgcgtcta cgaactgccg ataaacagag 16320
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380
aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500
tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggctctc aaacaggagg 16560
acggcccaa ggacgtcac aaggcgcac tgtccggcgt ttcgtggag ccgaacagc 16620
gaggccgagg ggtcgccggt atgctgctgc gggcgttgc ggcgggttta ttgctcgtga 16680
tgatcgctcg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgac 16740
ttaatatttc gctattctgg agcttggtt ttatttcggt ctaccgcctg ccgggcgggg 16800
tcgcggcgac ggtaggcgt gtgcagcgc tgatggctgt gttcatctc gccgctctgc 16860
taggtagccc gatacgattg atggcggctc tgggggctat ttgcggaact gcgggcgtgg 16920
cgctgttggt gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 16980
cgggggcgggt ttccatggcg ttcggaaccg tgctgacccg caagtggcaa cctcccgtgc 17040
ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 17100
ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggcctggcgt 17160
ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220

togaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280
cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag 17340
caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400
ttcctcagcg gctttatcca gcgatttcct attatgtcgg catagtcttc aagatcgaca 17460
gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580
ccgcgagatc atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc 17640
ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact 17700
gatgggctgc ctgtatcgag tggtgathtt gtgccgagct gccggtcggg gagctgttgg 17760
ctggctggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17820
acattgcgga cgtttttaat gtactggggg ggtttttctt ttcaccagtg agacgggcaa 17880
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggt 17940
ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18000
aaatcaaaag aatagcccga gatagggttg agtggtgttc cagtttggaa caagagtcca 18060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120
ccactacgtg aaccatcacc caaatcaagt tttttggggg cgagggtgccg taaagcacta 18180
aatcgggaacc ctaaaggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg 18240
gcgagaaagg aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg 18300
ggaagggcga tcgggtcggg cctcttcgct attacccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcgagctcg gtacccggg 18449

<210> 41
<211> 18449
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 41
gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60

cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120

cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180

cttctgatcc agtcgatcat ccgctgaagg cgcttttcgaa tctggttaag atccacgtct 240

tcgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttotca 300

gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360

gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg 420

atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc 540
aatattctca agtatctoga gtttgaactt attccctgtg aaccttttat tcaccaatga 600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720
tccgatagct ctgcaaaggg cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900
cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960
tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020
gtgaggaagt ggactcaa at cgacttcagc aacatctcct ggataaactt taagcctaaa 1080
ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140
gttgaccggg gcttggatct tcctatagaa tcatccttat tcgttgacct agctgattct 1200
ggagtgaccc agagggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260
aaatgtgacg aactcgtgag ctctgtacag tgaccgggtga ctctttctgg catgcggaga 1320
gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380
ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440
aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcatcggaga atatggagct 1500
tcatcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560
cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgatc tggtaaaaga 1620
ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680
aattggccca tccggcatct gtagggcgctc caaatatcgt gcctctcctg ctttgcccg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcaact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctgt ccgcccgggtg tgtcggcggg gttgacaagg tcgttgcgctc 1920

agccaacat ttgttgccat attttctctgc tctccccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc tcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctccggagggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcgtg gatatgtcct gcgggtaaatt agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgtc tgacattggg 2400

gaattcagcg agagcctgac ctattgcac tcccgccgtg cacagggtgt cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520

atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580

ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640

tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga ttccggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta cttcgagcgg aggcacccgg agcttgcagg atcgccgcgg 2940

ctccgggctg atatgctccg cattggctctt gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtcg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgcccg cgaagcgcg cgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctggtccgag ggcaaaggaa 3180
tagagtagat gccgaccgag ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300
tggttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagt cttcttagtg 3360
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600
atatttgtct ctaattgtac ttgacatgc tctcttctt tactctgata gcttgactat 3660
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780
gatggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa cacccaatac 3840
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900
cttggtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020
agaaagttaa aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080
ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140
ggagtttgaa attttgtcct ccaggagcaa aaaaaagttt ttttttatac atgtttgtac 4200
acaagaatag ttaccaatct gcttggtct tacgtgctgc aagtttatat cgttttcaat 4260

ttctttgtct ttacattttc tttgtccttt atctttcctc atttagtctt tgggagaatt 4320
aggaaaaggg agcggaaagg taagaaatgc ttgcgtatct tactaattcg gcaaacaatcc 4380
aatttggaac acagcagcct gtgcaacgct ctcgagatga cagtatcttt gattacactc 4440
taaattctga tgacccgacc aaaaagagcg aacaagaaa taatcttggt cattcgaata 4500
tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560
aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtcgaa 4620
gctgcagtca atcagcgtca agggccgccc cgttgaacta gcccgcgaca tcacgcggcc 4680
caaagtctgc ctgcatgctc agcgggtgctc gttagtctcg ctgcgagtgg cagcaccaca 4740
gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800
tgtagcactc cagcagcttg accgggctat cgagagcgt cgtgcccggc gcaaacggga 4860
gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920
cttcgccacc tacctgagat ttgccatgca catgaccgtg ggcggcgagc tgccatgggg 4980
tgaagtgggt ggcactctcc tcttggtggt tgggtggcgcg ctcggcatgg agatgtatgc 5040
ccgctatgca cacaaagcca tctggcatga gtcgcctctg ggctggctgc tgcacaagag 5100
ccaccacaca cctcgcactg gaccctttga agccaacgac ttgtttgcaa tcatcaatgg 5160
actgcccgcc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tgggggcggc 5220
ctgctttgga gcggggctgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280
tggcctgggt cacaggcgct tccccaccgg gccatcgct ggcctgccct acatgaagcg 5340
cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctgggggtat 5400
gttcttgggt ccacaggagc tgcagcacat tccaggtgcg gcggaggagg tggagcgact 5460
ggctctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580

cgcccgtagc ctaggtgatc aaagactgtt gggcttgtct ctgaagcttg taggaaaggc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760

gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaaga ttgctctgtt 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880

gattaagagt ttttttgttg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

aaagagacaa cacatacaat ataaaatcat tgtcttttaa tgtaattta gtagagtga 6000

aagcctgcat tttttttgta cgcataaaca atgaattcac cccgcttctg gtttttaa 6060

aattatgtca aactagggaa aattcttttt tttctcttcg ttcttttttt ggcttggtgt 6120

ggagtcacag gcttgtcttc agattgatag aggttgata cactcaacag agcaatcttg 6180

gcacgttcgc ttccttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt 6240

acagatccaa ttgcagcata tacttgga gacgctcgta caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcctacaagc 6360

ttcagagaca agcccaacag tctttgatca cctaggctac gggcgttgcc ttgctgtatt 6420

tgttctgttt cttcttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480

gagccatact acagcataaa aggatacggtt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac aaaaaaaaaa aaatttaaga tgagtaggac ttccattctc tcaaaaattt 6600

tattcaatcc ataaatgaat tatttttgga caaaaaagaa agattatgcc tgattttctc 6660

tatttttttt ttttttacia ctccaccaat actttctagc ccagcttggc gtaatcatgg 6720

tcatagctgt ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc 6780

ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg 6840

ttgcgctcac tgcccgcctt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900

ggccaacgcg cggggagagg cggtttgctg attgggcca agacaaaagg gcgacattca 6960

accgattgag ggaggggaagg taaatattga cggaattat tcattaaagg tgaattatca 7020

ccgtcaccga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140

gcgacagaat caagtgtgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg 7200

gtcatagccc ccttattagc gtttgccatc ttttcataat caaatcacc ggaaccagag 7260

ccaccaccgg aaccgcctcc ctcagagccg ccaccctcag aaccgccacc ctcagagcca 7320

ccaccctcag agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc 7380

ccgatctagt aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat 7440

tttgttttct atcgcgctatt aaatgtataa ttgcgggact ctaatcataa aaacccatct 7500

cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560

ttatatgata atcatcgcaa gaccggcaac aggattcaat cttagaaac tttattgcca 7620

aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680

acgcagcaag atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt 7740

gatgtggacg ccgggcccga tcatattgtc gctcaggatc gtggcgttgt gcttgtcggc 7800

cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccggtggc 7860

gaagaactcc agcatgagat cccgcgctg gaggatcatc cagccggcgt cccggaaaac 7920

gattccgaag cccaaccttt catagaaggc ggcgggtggaa tcgaaatctc gtgatggcag 7980

gttgggcgtc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040

agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100
aagcggtcag cccattcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg 8160
tcctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220
ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg 8280
tcgggcatgc gcgccttgag cctggcgaa agttcggctg gcgcgagccc ctgatgctct 8340
tcgtccagat catcctgac gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg 8400
cgatgtttcg cttgggtggc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460
attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520
tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc 8580
acagctgcgc aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctgctcctgc 8640
agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct 8700
gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg 8760
aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820
cgaaacgac cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880
ggataccgag gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940
gctgatagt accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000
agctcattaa actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca 9060
gttccaaacg taaaacggct tgtcccgcgt catcggcggg ggtcataacg tgactccctt 9120
aattctccgc tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 9180
acaggatata ttggcgggta aacctaaagag aaaagagcgt ttattagaat aatcggatat 9240
ttaaaggggc gtgaaaagggt ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300

gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgcccgc ccgaaacgat 9360
ccgacagcgc gccagcaca ggtgcgagg caaattgcac caacgcatac agcgccagca 9420
gaatgccata gtgggcggtg acgtcgttcg agtgaaccag atcgcgagg aggcccggca 9480
gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgctc agaattacga 9540
tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600
cgcgcggtt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660
gtcaagcatg acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa 9720
cgaggtcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780
gccggcgctt tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc 9840
catgctggcg gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt 9900
tctgatcggg aatgcccgca gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg 9960
catccatgcc ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg 10020
cttcctctgc gaggcgggtt ttccggccgg ggacgccgctc aatgcgctga tgacaatcag 10080
ctacttcact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140
cggcggcacc gttgaacagg ctccgctctc gccgctgttg cgggccgcga tagacgcctt 10200
cgacgaagcc ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt 10260
ggcgaaaagg aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc 10320
aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380
cccccttcc accgcgtcag acgcccgtag cagcccgcta cgggcttttt catgccctgc 10440
cctagcgtcc aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500
gcttctctgc tcaactgact gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 10560

cactcaaagg cggtaatagc gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620

tgagcaaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680

cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740

aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 10800

cctgttccga ccctgccgt taccggatac ctgtccgcct ttctcccttc gggaagcgtg 10860

gcgctttttc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc 10920

atcctttttc gcacgatata caggattttg ccaaagggtt cgtgtagact ttccttggtg 10980

tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc 11040

cttcttctact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100

ctggccggct accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160

aaccaggaag ggcagcccac ctatcaagggt gtactgcctt ccagacgaac gaagagcgat 11220

tgaggaaaag ggggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280

gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340

tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 11400

ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460

cgagcttggc aaggtcatga tgggcgtggt ccgcccaggg gcagagccat gactttttta 11520

gccgctaaaa cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580

agaagagcga cttcgcgag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc 11640

gcctttgcga cgctcaccgg gctgggtgcc ctgcgcgtg ggctggcggc cgtctatggc 11700

cctgcaaacg cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt 11760

tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820

tgaggggccc actcaccg cgcggcgttg acagatgagg ggcaggctcg atttcggccc 11880

gcgacgtgga gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc 11940

ccacagatga tgtggacaag cctggggata agtggcctgc ggtattgaca cttgaggggc 12000

gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060

tgagggggcg acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc 12120

aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 12180

atatattataa accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacggcg 12240

aaggggggtg ccccccttc tcgaaccctc ccggcccgt aacgcgggccc tcccatcccc 12300

ccaggggctg cgcccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt 12360

ccttgccatt gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc 12420

cggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480

tgagggcggc ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga 12540

cttcatggcg gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcggggtgc 12600

cgtgctcgtg ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt 12660

ataccgaggt atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat 12720

ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ctttgaatat 12780

attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840

tttcaggggg caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaaagca 12900

taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960

ctatcataat tgggtaatat ctccaactta ttgatagtgt tttatgttca gataatgccc 13020

gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg 13080

tgccaggtgc tgcctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140
gattacgtgc agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca 13200
tatcaccacg tcaaagggtg acagcaggct cataagacgc cccagcgtcg ccatagtgcg 13260
ttcaccgaat acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca 13320
gcgctggcgc gatttagccc cgacatagcc ccactgttcg tccatttccg cgcagacgat 13380
gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggcttgagtt ttttaagtga 13440
cgtaaaatcg tgttgaggcc aacgccata atgcgggctg ttgccggca tccaacgcca 13500
ttcatggcca tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac 13560
tgcagttgcc atgttttacg gcagtgcgag cagagatagc gctgatgtcc ggcggtgctt 13620
ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680
gccactggag cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc 13740
cataattgtg gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaaac 13800
aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860
gagttcgtct tggtataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920
ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980
accgctgcgt aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040
aaaatgaaaa cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg 14100
tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc 14160
tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc 14220
tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280
agtgcacacg gctctttcac tccatcgaca tatcggattg tccctatagc aatagcttag 14340

acagccgctt agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg 14400
aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaga 14460
cggaaaagcc cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520
ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580
agtggatatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt 14640
atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700
atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760
caggagcgca ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc 14820
aagtatttgg gcaaggggtc gctgggtattc gtgcagggca agattcggaa taccaagtac 14880
gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940
gacaccaagg caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtgagtc 15000
ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060
gaactgatcg acgcggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc 15120
atgcgtgcgc cccgcgaaac cttccagtcc gtcggctcga tgggtccagca agctacggcc 15180
aagatcgagc gcgacagcgt gcaactggct cccctgccc tgcccgcgc atcggcgcgc 15240
gtggagcgtt cgcgtcgtct cgaacaggag gcggcaggtt tggcgaagtc gatgaccatc 15300
gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa 15360
caggtcagcg aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420
atgcagcttt ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaaac 15480
gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540
caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag 15600

ctgcgggccc acgatgacga actggtgtgg cagcaggtgt tggagtacgc gaagcgcacc 15660
cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggtcg 15720
atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780
atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgcgctgct gcaccgcttc 15840
cgcgtcctgg accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc 15900
gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960
tcgccgacgg cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgctc 16020
aagctggaaa ccttcgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080
gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctgggtgga acacgcctgg 16140
gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200
ggttcagcag ccagcgtttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260
tgcttcgctc agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag 16320
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380
aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500
tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 16560
acggcccaa ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc 16620
gaggccgagg ggtcgccggt atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga 16680
tgatcgctcc acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac 16740
ttaatatttc gctattctgg agcttggtgt ttatttcgggt ctaccgcctg ccgggcgggg 16800
tcgcggcgac ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgctctgc 16860

taggtagccc gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg 16920
cgctgttggg gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 16980
cgggggcggt ttccatggcg ttcggaaccg tgctgacctg caagtggcaa cctcccgtgc 17040
ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 17100
ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggcctggcgt 17160
ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220
tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280
cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag 17340
caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400
ttcctcagcg gctttatcca gcgatttcct attatgtcgg catagtcttc aagatcgaca 17460
gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580
ccgcgagatc atccgtgttt caaacccggc agcttagttg ccgttcttcc gaatagcatc 17640
ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact 17700
gatgggctgc ctgtatcgag tgggtgatttt gtgccgagct gccggtcggg gagctgttgg 17760
ctggctgggtg gcaggatata ttgtgggtgta aacaaattga cgcttagaca acttaataac 17820
acattgcgga cgtttttaat gtactggggg ggtttttctt ttcaccagtg agacgggcaa 17880
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg 17940
ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18000
aaatcaaaag aatagccccg gataggggtg agtggtgttc cagtttgga caagagtcca 18060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120

ccactacgtg aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta 18180
aatcggaacc ctaaagggag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg 18240
gcgagaaagg aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg 18300
ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcgagctcg gtacccggg 18449

<210> 42
<211> 17593
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 42
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttcccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaaggt ttccgcccgt	480
ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaa cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcgcccc	660
tcggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgca cggccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccg gcagtgaggg cggcggcctg	840
ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140
agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc	1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga	1260
ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta	1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc	1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctggtgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagtgc gtcttgttat 2220

aattagcttc ttggggatat tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccatct aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatcttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac caccgcgaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactgggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgetgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgcac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccg gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcattgcaa acgctagggc cttgtgggggt cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gtcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gtcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860

catcgacggc gagatcattg ggctgtcgggt cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggctg tgccggcggg ttattgtctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatctt catcctcggc gcaacttaata tttcgctatt 5100

ctggagcttg ttgtttatct cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtccctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgttgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagt gcaacctccc gtgcctctgc tcacctttac 5400
cgccctggcaa ctggcgggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgccgtcccc gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300
aaatcctgtt tgatgggtgg tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900
gtcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgtga aggcgctttc gaatctggtt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttggg gcatagaggg 7320
tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440
tcgaaataca tccgatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctogagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920
tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga cccagagggt catgacttga gcctaaaadc cgccgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggtc agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580
cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggctcgttc gtcagtcctc catttggttc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttggtacat ccatactcca tccttcccat cccttattcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagtctgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccatttcgg 9360
accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctcg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgtc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacggcagc tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccacaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctggtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ttttcgagtt 10800
tttttttttt ttctttgtga aggatttatt gttattggta tccatttttt attggaagac 10860
aagataagtt aatattgatt ttgcttaaag attaaaagga aatcagaaaa cgacaataaa 10920
aaatgtaacg gacaaactat ggtgtcgatt ataagtctaa atccttaaaa aatgacaacg 10980
agttgctttc ctctgaaaac aattcttttg tctttgcaag aaaggtttct tttttgtttg 11040
cttgcattac ttaaacaatca aatcaaataa aaggaataaa gcagatttga gggcgaataa 11100
ggattttctg gtcaacaaga tgtgagtgac acctaaaggaa ctaaagcca ttcatttggt 11160
ttaaaacgac atcaaagatt gatgatcaac aggattgaga gagagaaaaa gaactcgtgt 11220
catttatttc tgttgactga aattttatat ttagaaaaaa tgtcaaactc atagctttag 11280
ctatattaca taacatttga aataataata ataaaaaaag acacattaga gacacttttc 11340
aaactctaaa taactgtcta taaacacaaa gaaaacaaag acctctataa caacttatta 11400
gatttttctc gtacttttgt ctaaagatga tgtattcttg ttatcccaca cttctttcat 11460

ttgttcttga tgctactaaa tatacaaaat ttcttttttg caagagatat tattccaaaa 11520
attttcaaaa agaaatTTTT ttcacaatag cagttgatcg tgtaacccaa agaggttctt 11580
tgttattttg cacttccgct ttgcggtgat gcatattcaa agtaatatat ggaataaaca 11640
acgtgtttta gcatgaaaga aaggaaacaa aggccgcttt gaacaaatgc ataatatctt 11700
agacaaaaat gatctaaagc aagcagtaaa tcaaacaaga aacattgctg attcgcgtta 11760
gaaaacgata aaagtctaata aagccactaa gtatacttca atgaactttt tgtatgctta 11820
tggtccaatc agaccaataa tttgtgacca ttcctgaggt ggctttggtg atgcggaaac 11880
agaaaaaaat tttctcacca atcgatttaa aaaacaattt ctgctttgaa caaaactttt 11940
ttttttctct ttaatcatta actttatcaa gtatgtacct accctcaaag tcttcaactca 12000
agcacatta tgctaacatt gttccacctt ctctttagaa atgttggtgga tttggaatgc 12060
cctgatcggt ttcgttaccg tgattggcat ggaagtgatt gctgcactgg cacacaaata 12120
catcatgcac ggctgggggtt ggggatggca tctttcacat catgaaccgc gtaaagggtgc 12180
gtttgaagtt aacgatcttt atgccgtggt ttttgctgca ttatcgatcc tgctgattta 12240
tctgggcagt acaggaatgt ggccgctcca gtggattggc gcaggatga cggcgatatg 12300
attactctat tttatggtgc acgacgggct ggtgcatcaa cgttggccat tccgctatat 12360
tccacgcaag ggctacctca aacggttgta tatggcgcac cgtatgcatc acgccgtcag 12420
gggcaaagaa ggttgtgttt cttttggctt cctctatgcg ccgcccctgt caaaacttca 12480
ggcgacgctc cgggaaagac atggcgctag agcgggctg gccagagatg cgcagggcgg 12540
ggaggatgag cccgcatccg ggaagtaagg gcctgaccag aggcggccag cagcagcggt 12600
aatttttcgg gcgtggctgt tgactgccgc tgatcccaaa gcttggcgta atcatggtca 12660
tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc cacacaacat acgagccgga 12720

agcataaagt gtaaagcctg ggggtgcctaa tgagtgaagt aactcacatt aattgcgttg 12780
cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc 12840
caacgcgcgg ggagaggcgg tttgcgtatt gggccaaaga caaaagggcg acattcaacc 12900
gattgagggg ggggaaggtaa atattgacgg aaattattca ttaaagggtga attatcacgg 12960
tcaccgactt gagccatttg ggaattagag ccagcaaaat caccagtagc accattacca 13020
ttagcaaggc cggaaacgtc accaatgaaa ccatcgatag cagcacgta atcagtagcg 13080
acagaatcaa gtttgccttt agcgtcagac tgtagcgcgt tttcatcggc attttcggtc 13140
atagccccct tattagcgtt tgccatcttt tcataatcaa aatcacggga accagagcca 13200
ccaccggaac cgcctccctc agagccgcca ccctcagaac cgccaccctc agagccacca 13260
ccctcagagc cgccaccaga accaccacca gagccgccgc cagcattgac aggaggcccc 13320
atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgcgc gctatatttt 13380
gttttctatc gcgtattaaa tgtataattg cgggactcta atcataaaaa cccatctcat 13440
aaataacgtc atgcattaca tgttaattat tacatgctta acgtaattca acagaaatta 13500
tatgataatc atcgcaagac cggcaacagg attcaatctt aagaaacttt attgccaaat 13560
gtttgaacga tcggggatca tccgggtctg tggcggaac tccacgaaaa tatccgaacg 13620
cagcaagata tcgcggtgca tctcgggtct gcctgggcag tcgccgccga cgccgttgat 13680
gtggacgccg ggcccgatca tattgtcgct caggatcgtg gcgttggtgt tgtcggccgt 13740
tgctgtcgta atgatatcgg caccttcgac cgcctgttcc gcagagatcc cgtgggcgaa 13800
gaactccagc atgagatccc cgcgctggag gatcatccag ccggcgctcc ggaaaacgat 13860
tccgaagccc aacctttcat agaaggcggc ggtggaatcg aaatctcgtg atggcaggtt 13920
gggcgtcgct tggtcgggtca tttcgaacct cagagtcccc ctcagaagaa ctcgtaaga 13980

aggcgataga aggcgatgcg ctgcgaatcg ggagcggcga taccgtaaag cacgaggaag 14040
cggtcagccc attcgccgcc aagctcttca gcaatatcac gggtagccaa cgctatgtcc 14100
tgatagcggg cgcacacacc cagccggcca cagtcgatga atccagaaaa gcggccattt 14160
tccaccatga tattcggcaa gcaggcatcg ccatgggtca cgacgagatc atcgccgtcg 14220
ggcatgcgcg ccttgagcct ggccaacagt tcgggtggcg cgagcccctg atgctcttcg 14280
tccagatcat cctgatcgac aagaccgggt tccatccgag tacgtgctcg ctcgatgcga 14340
tgtttcgctt ggtggtcgaa tgggcaggta gccggatcaa gcgtatgcag ccgccgcatt 14400
gcatcagcca tgatggatac tttctcggca ggagcaagggt gagatgacag gagatcctgc 14460
cccggcactt cgccaatag cagccagtcc cttcccgctt cagtgacaac gtcgagcaca 14520
gctgcgcaag gaacgcccgt cgtggccagc cacgatagcc gcgctgcctc gtccctgcagt 14580
tcattcaggg caccggacag gtcggctctt acaaaaagaa cggggcgccc ctgcgctgac 14640
agccggaaca cggcggcatc agagcagccg attgtctggt gtgcccagtc atagccgaat 14700
agcctctcca cccaagcggc cggagaacct gcgtgcaatc catcttggtc aatcatgcga 14760
aacgatccag atccgggtgca gattatttgg attgagagtg aatatgagac tctaattgga 14820
taccgagggg aatttatgga acgtcagtgg agcatttttg acaagaaata tttgctagct 14880
gatagtgacc ttaggcgact tttgaacgcg caataatggt ttctgacgta tgtgcttagc 14940
tcattaaact ccagaaaccc gcggctgagt ggctccttca acgttgcggt tctgtcagtt 15000
ccaaacgtaa aacggcttgt ccgcgctcat cggcgggggt cataacgtga ctcccttaat 15060
tctccgctca tgatcagatt gtcgtttccc gccttcagtt taaactatca gtgtttgaca 15120
ggatatattg gcgggtaaac ctaagagaaa agagcgttta ttagaataat cggatattta 15180
aaagggcgtg aaaaggttta tccgttcgtc catttgatg tgcatgccaa ccacagggtt 15240

ccccagatct ggcgccggcc agcgagacga gcaagattgg cgcgccccg aaacgatccg 15300
acagcgcgcc cagcacaggt gcgcaggcaa attgcaccaa cgcatacagc gccagcagaa 15360
tgccatagtg ggcggtgacg tcgttcgagt gaaccagatc gcgcaggagg cccggcagca 15420
ccggcataat caggccgatg ccgacagcgt cgagcgcgac agtgctcaga attacgatca 15480
ggggtatgtt gggtttcacg tctggcctcc ggaccagcct ccgctgggtcc gattgaacgc 15540
gcggattctt tatcactgat aagttggtgg acatattatg tttatcagtg ataaagtgtc 15600
aagcatgaca aagttgcagc cgaatacagt gatccgtgcc gccctggacc tgttgaacga 15660
ggtcggcgta gacggtctga cgacacgcaa actggcgga cggttggggg ttcagcagcc 15720
ggcgctttac tggcacttca ggaacaagcg ggcgctgctc gacgcactgg ccgaagccat 15780
gctggcgagg aatcatacgc attcggtgcc gagagccgac gacgactggc gtcattttct 15840
gatcgggaat gcccgagct tcaggcaggc gctgctcgcc taccgcatg gcgcgcgcat 15900
ccatgccggc acgcgaccgg gcgcaccgca gatggaaacg gccgacgcgc agcttcgctt 15960
cctctgcgag gcgggttttt cggccgggga cgccgtcaat gcgctgatga caatcagcta 16020
cttcactgtt ggggccgtgc ttgaggagca ggccggcgac agcgatgccg gcgagcgcg 16080
cggcaccgtt gaacaggctc cgctctcgcc gctgttgcg gccgcgatag acgccttcga 16140
cgaagccggt ccggacgcag cgttcgagca gggactcgcg gtgattgtcg atggattggc 16200
gaaaaggagg ctcgttgtca ggaacgttga aggaccgaga aagggtgacg attgatcagg 16260
accgctgccg gagcgcaacc cactcactac agcagagcca tgtagacaac atcccctccc 16320
cctttccacc gcgtcagacg cccgtagcag cccgctacgg gctttttcat gccctgcct 16380
agcgccaag cctcacggcc gcgctcgcc tctctggcgg cctctggcg ctcttcgct 16440
tctcgtca ctgactcgt gcgctcggtc gttcggctgc ggcgagcgg atcagctcac 16500

tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga 16560
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cggtgctggc gtttttccat 16620
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 16680
ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 16740
gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 16800
cttttccgct gcataaccct gcttcgggggt cattatagcg attttttcgg tatatccatc 16860
ctttttcgca cgatatacag gattttgcc aagggttcgt gtagactttc cttggtgtat 16920
ccaacggcgt cagccgggca ggatagggtga agtaggccca cccgcgagcg ggtgttcctt 16980
cttcactgtc ccttatcgc acctggcggt gctcaacggg aatcctgctc tgcgaggctg 17040
gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa ccaagccaac 17100
caggaagggc agcccaacct tcaaggtgta ctgccttcca gacgaacgaa gagcgattga 17160
ggaaaaggcg gcggcgggcg gcatgagcct gtcggcctac ctgctggccg tcggccaggg 17220
ctacaaaatc acgggcgctg tggactatga gcacgtccgc gagctggccc gcatcaatgg 17280
cgacctgggc cgcctgggcg gcctgctgaa actctggctc accgacgacc cgcgcacggc 17340
gcggttcggt gatgccacga tcctcgccct gctggcgaag atcgaagaga agcaggacga 17400
gcttggcaag gtcattgatg gcgtgggtccg cccgagggca gagccatgac ttttttagcc 17460
gctaaaacgg ccgggggggtg cgctgattg ccaagcacgt ccccatgcgc tccatcaaga 17520
agagcgactt cgcgagctg gtgaagtaca tcaccgacga gcaaggcaag accgagcgcc 17580
tttgcgacgc tca 17593

<210> 43

<211> 16954

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 43

ccgggctggt tgcctcgcc gctgggctgg cggcgtcta tggcctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gggccgccg gcgttggtga tacctcgccg	120
aaaacttggt cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt	480
ttttcgcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600

cttctcgaac cctccccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc 660
tcggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720
atcggggcag taacgggatg ggcgatcagc ccgagcgca cgcccgaag cattgacgtg 780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840
ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900
gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140
agataatata tcttttatat agaagatata gccgtatgta aggatttcag ggggcaaggc 1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260
ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtocca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttcggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta tttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcgagtgca tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatatatt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcggggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcgggggca atcccgaag 3360
gaggggtgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggtcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaata accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
ccctgttcac cgcgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggccccgac 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggccttg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatcgaa acgctagggc cttgtggggt cagtccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccga cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggctg tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata tttcgtatt 5100

ctggagcttg ttgtttatct cggctctaccg cctgcggggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatttgctg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgcctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggctgat cagccgggga tgcattcaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacgggtaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta cctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgccgtcccg gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctctttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320
tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcg catgagggtt 7440
tcgaaatata tccggatgtc gaaggcttgg ggcacctgag ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaa gggcgttgca caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaattc 7920
tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcggtga 7980
cctagctgat tctggagtga cccagagggg catgacttga gcctaaaatt cgcgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttcg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccc gtgtgtcggc ggggttgaca 8700
aggtegttg gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccg 8940
ttgagcagac atcaccatgc ctgaactcac cggacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgt gttctgcagc cggtcgcgga 9300
ggccatggat gcgacgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360
accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcggtccg tcgcgagggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggcgcg ataacagcgg tcattgactg 9600
gagcgagggc atgttcgggg attcccaata cgagggtgcc aacatcttct tctggaggcc 9660
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgtc agtccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttgttt agaggtaatc cttcttttcta gctagaagtc ctctgtgtact 10740
gtgtaagcgc ccactccaca tctccacteg acctgcaggc atgcaagctt gagattaaaa 10800
tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt cttttttataa 10860
atacatgcct gacttttctt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160
cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac tctaaatctc gatgaccoga ccaaaaagag cgaacaaaga aataatcttg 11280
tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340
tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400
tatgttgtgg atttggaatg ccctgatcgt tttcgttacc gtgattggca tggaagtgat 11460
tgctgcactg gcacacaaat acatcatgca cggctgggggt tggggatggc atctttcaca 11520
tcatgaaccg cgtaaagggt cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580
attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640
cgcaggtatg acggcgtatg gattactcta ttttatgggt cacgacgggc tgggtgcatca 11700
acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760
ccgtatgcat cagccgtca ggggcaaaga aggttgtgtt tcttttggct tcctctatgc 11820
gccgcccctg tcaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcgggcgc 11880
tgccagagat gcgcagggcg gggaggatga gccgcgcatcc gggaagtaag ggcctgacca 11940

gaggcggcca gcagcagcgt taatTTTTcg ggcgtggctg ttgactgccg ctgatcccaa 12000
agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gtcacaatt 12060
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 12120
taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc 12180
cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggccaaag 12240
acaaaagggc gacattcaac cgattgaggg agggaaggta aatattgacg gaaattattc 12300
attaaaggtg aattatcacc gtcaccgact tgagccattt gggaattaga gccagcaaaa 12360
tcaccagtag caccattacc attagcaagg ccggaacgt caccaatgaa accatcgata 12420
gcagcacctg aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12480
ttttcatcgg cattttcggg catagcccc ttattagcgt ttgccatctt ttcataatca 12540
aaatcacccg aaccagagcc accaccggaa ccgcctccct cagagccgcc accctcagaa 12600
ccgccaccct cagagccacc accctcagag ccgccaccag aaccaccacc agagccgcgcg 12660
ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataattttatc 12720
ctagtttgcg cgctatatTT tgTTTTctat cgcgtattaa atgtataatt gcgggactct 12780
aatcataaaa acccatctca taaataacgt catgcattac atgttaatta ttacatgctt 12840
aacgtaattc aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 12900
taagaaactt tattgccaaa tgtttgaacg atcggggatc atccgggtct gtggcgggaa 12960
ctccacgaaa atatccgaac gcagcaagat atcgcgggtgc atctcgggtct tgcctgggca 13020
gtcgccgcgc acgcggttga tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt 13080
ggcgttgtgc ttgtcggccg ttgctgtcgt aatgatatcg gcaccttoga ccgcctgttc 13140
cgcagagatc ccgtgggcga agaactccag catgagatcc ccgcgctgga ggatcatcca 13200

gccggcggtcc cggaaaacga ttccgaagcc caacctttca tagaaggcgg cggtggaatc 13260
gaaatctcgt gatggcaggt tgggcgtcgc ttggtcggtc atttcgaacc ccagagtccc 13320
gctcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc gggagcggcg 13380
ataccgtaaa gcacgaggaa gcggtcagcc cattcgccgc caagctcttc agcaatatca 13440
cgggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggcc acagtcgatg 13500
aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc 13560
acgacgagat catcgccgtc gggcatgcgc gccttgagcc tggcgaacag ttcggctggc 13620
gcgagcccct gatgctcttc gtccagatca tcctgatcga caagaccggc ttccatccga 13680
gtacgtgctc gctcgatgcg atgtttcgct tggtggtcga atgggcaggt agccggatca 13740
agcgtatgca gccgccgat tgcattcagcc atgatggata ctttctcggc aggagcaagg 13800
tgagatgaca ggagatcctg ccccggcact tcgccaata gcagccagtc ccttcccgtc 13860
tcagtgaaa cgtcgagcac agctgcgcaa ggaacgcccg tcgtggccag ccacgatagc 13920
cgcgctgcct cgtcctgcag ttcatlcagg gcaccggaca ggtcggctctt gacaaaaaga 13980
accgggcgcc cctgcgctga cagccggaac acggcgccat cagagcagcc gattgtctgt 14040
tgtgcccagt catagccgaa tagcctctcc acccaagcgg cgggagaacc tgcgtgcaat 14100
ccatcttgtt caatcatgcg aaacgatcca gatccggtgc agattatttg gattgagagt 14160
gaatatgaga ctctaattgg ataccgaggg gaatttatgg aacgtcagtg gagcattttt 14220
gacaagaaat atttgctagc tgatagtac cttaggcgac ttttgaacgc gcaataatgg 14280
tttctgacgt atgtgcttag ctcatataac tccagaaacc cgcggtgag tggctccttc 14340
aacgttgccg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg 14400
tcataacgtg actcccttaa ttctccgctc atgatcagat tgcgttttcc cgccttcagt 14460

ttaaactatc, agtgtttgac aggatatatt ggcgggtaaa cctaagagaa aagagcgttt 14520
attagaataa tcggatatTTT aaaagggcgt gaaaaggttt atccgttcgt ccatttgtat 14580
gtgcatgccca accacagggT tccccagatc tggcgccggc cagcgagacg agcaagattg 14640
gccgccgccc gaaacgatcc gacagcgcg cagcacagg tgcgcaggca aattgcacca 14700
acgcatacag cgccagcaga atgccatagt gggcggtgac gtcgttcgag tgaaccagat 14760
cgcgcaggag gcccggcagc accggcataa tcaggccgat gccgacagcg tcgagcgcg 14820
cagtgtcag aattacgac aggggtatgt tgggtttcac gtctggcctc cggaccagcc 14880
tccgctggtc cgattgaacg cgcggttct ttatcactga taagttggtg gacatattat 14940
gtttatcagt gataaagtgt caagcatgac aaagttgcag ccgaatacag tgatccgtgc 15000
cgccctggac ctgttgaacg aggtcggcgt agacggtctg acgacacgca aactggcgga 15060
acggttgggg gttcagcagc cggcgcttta ctggcacttc aggaacaagc gggcgctgct 15120
cgacgcactg gccgaagcca tgctggcgga gaatcatac cattcggtgc cgagagccga 15180
cgacgactgg cgctcatttc tgatcgggaa tgcccgcagc ttcaggcagg cgctgctcgc 15240
ctaccgcgat ggcgcgcga tccatgcgg cagcgaccg ggcgcaccgc agatggaaac 15300
ggccgacgcg cagcttcgct tcctctgcga ggcgggtttt tcggccgggg acgccgtcaa 15360
tgcgctgatg acaatcagct acttcactgt tggggccgtg cttgaggagc aggccggcga 15420
cagcgatgcc ggcgagcgcg gcggcaccgt tgaacaggct ccgctctcgc cgctgttgcg 15480
ggccgcgata gacgccttcg acgaagccgg tccggacgca gcgttcgagc agggactcgc 15540
ggtgattgtc gatggattgg cgaaaaggag gctcgttgtc aggaacgttg aaggaccgag 15600
aaaggtgac gattgatcag gaccgctgcc ggagcgcaac cactcacta cagcagagcc 15660
atgtagacaa catccccctc ccctttccac cgcgtcagac gcccgtagca gcccgctacg 15720

ggctttttca tgccctgccc tagcgtccaa gcctcacggc cgcgctcggc ctctctggcg 15780
gccttctggc gctcttcgc ttcctcgctc actgactcgc tgcgctcggc cgttcggctg 15840
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat 15900
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 15960
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc 16020
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 16080
agctccctcg tgcgctctcc tgttcgcacc ctgccgctta ccggatacct gtccgccttt 16140
ctcccttcgg gaagcgtggc gcttttcgc tgcataacc tgcttcgggg tcattatagc 16200
gattttttcg gtatatccat ctttttcgc acgatataca ggattttgcc aaagggttcg 16260
tgtagacttt ccttggtgta tccaacggcg tcagccgggc aggatagggtg aagtaggccc 16320
accgcgagc ggggtgttct tcttactgt cccttattcg cacctggcgg tgctcaacgg 16380
gaatcctgct ctgcgaggct ggccggctac cgccggcgta acagatgagg gcaagcggat 16440
ggctgatgaa accaagccaa ccaggaaggc cagcccacct atcaagggtg actgccttcc 16500
agacgaacga agagcgattg aggaaaaggc ggccggcgcc ggcatgagcc tgcggccta 16560
cctgctggcc gtcggccagg gctacaaaat cacgggcgct gtggactatg agcacgtccg 16620
cgagctggcc cgcacatg ggcacctggc ccgcctgggc ggctgctga aactctggct 16680
caccgacgac ccgcgcacgg cgcgggttcg tgatgccacg atcctcgccc tgctggcgaa 16740
gatcgaagag aagcaggacg agcttgccaa ggtcatgatg ggctggtcc gcccgagggc 16800
agagccatga ctttttttagc cgctaaaacg gccggggggg gcgcgtgatt gccaaagcacg 16860
tccccatgcg ctccatcaag aagagcgact tcgaggagct ggtgaagtac atcaccgacg 16920
agcaaggcaa gaccgagcgc ctttgcgacg ctca 16954

<210> 44
<211> 16954
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 44
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggacagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcgggcga aacgcctgat ttacgcgag tttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatatth ataaacctg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctcccggcc cgctaacgcg ggccctcccat cccccaggg gctgcgcccc 660

tcggccgcga acggcctcac cccaaaaatg gcagcgcctg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgcga cgcccggag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggccggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatactata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgctt aaaccagga caataacct atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgcgcctc aattogetgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cagctcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgcgt gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaaggtata taagctgggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggttggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc ttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360
gagggtgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
ccctgttcac cgcgcgaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggt gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcg cctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttattt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtccctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgttgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgctggcaa ctggcggccg gaggacttct gctcgttcca gtagcttttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt totcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta cctccgcga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgccgtccc gactgatggg ctgcctgtat 6060
cgagtgggta ttttgtgccg agctgccggc cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300
aaatcctggt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggt 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcgggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcc 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggtat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaatata tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gtcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaata 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaata cgccgcctcc 8040

accattttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaatctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttcgg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctacccgcg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctctgtcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccatttcg 9360
accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcggtccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggctgcc aacatcttct tctggaggcc 9660
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc a gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtcaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttectac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800
taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860
tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920
ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca aagatactgt 10980
catctcgaga gcgttgaca ggctgctggt tgccaaattg gatgtttgcc gaattagtaa 11040
aatacgcaag catttcttac ctttccgctc ctttttcta attctcccaa agactaaatg 11100
aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga tataaacttg 11160
cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaacatg tataaaaaaa 11220
aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat tgcttgata 11280
tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta tttataaaaa 11340
gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct attttaatct 11400
catgttgtgg atttggaatg ccctgatcgt tttcgttacc gtgattggca tggaagtgat 11460
tgctgcactg gcacacaaat acatcatgca cggctggggg tggggatggc atctttcaca 11520
tcatgaaccg cgtaaagggt cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580
attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640
cgcaggtatg acggcgtatg gattactcta ttttatgggt cacgacgggc tgggtgcatca 11700
acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760
cogtatgcat cagccgtca ggggcaaaga aggttgtggt tcttttggct tcctctatgc 11820

gccgcccctg tcaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcggggcg 11880
tgccagagat gcgcagggcg gggaggatga gcccgcatcc gggaagtaag ggcctgacca 11940
gaggcggcca gcagcagcgt taatTTTTcg ggcgtggctg ttgactgccg ctgatcccaa 12000
agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gtcacaatt 12060
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 12120
taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc 12180
cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggccaaaag 12240
acaaaagggc gacattcaac cgattgaggg agggaaggta aatattgacg gaaattattc 12300
attaaagggtg aattatcacc gtcaccgact tgagccattt gggaattaga gccagcaaaa 12360
tcaccagtag caccattacc attagcaagg ccggaacgt caccaatgaa accatcgata 12420
gcagcaccgt aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12480
ttttcatcgg cattttcggc catagcccc ttattagcgt ttgccatctt ttcataatca 12540
aatcaccgg aaccagagcc accaccggaa ccgcctccct cagagccgcc accctcagaa 12600
ccgccaccct cagagccacc accctcagag ccgccaccag aaccaccacc agagccgccg 12660
ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataattttatc 12720
ctagtttgcg cgctatattt tgttttctat cgcgtattaa atgtataatt gcgggactct 12780
aatcataaaa acccatctca taaataacgt catgcattac atgttaatta ttacatgctt 12840
aacgtaattc aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 12900
taagaaactt tattgccaaa tgtttgaacg atcggggatc atccgggtct gtggcgggaa 12960
ctccacgaaa atatccgaac gcagcaagat atcgcgggtgc atctcggtct tgcctgggca 13020
gtgcgcccg acgccgttga tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt 13080

ggcgttgtgc ttgtcggccg ttgctgtcgt aatgatatcg gcaccttcga ccgcctgttc 13140
cgagagatc ccgtggggcga agaactccag catgagatcc ccgcgctgga ggatcatcca 13200
gccggcgctcc cggaaaacga ttccgaagcc caacctttca tagaaggcgg cggtggaatc 13260
gaaatctcgt gatggcaggt tgggcgtcgc ttggtcggtc atttcgaacc ccagagtccc 13320
gctcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc gggagcggcg 13380
ataccgtaaa gcacgaggaa gcggtcagcc cattcgccgc caagctcttc agcaatatca 13440
cgggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggcc acagtcgatg 13500
aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc 13560
acgacgagat catcgccgtc gggcatgcgc gccttgagcc tggcgaacag ttcggtctggc 13620
gcgagccctt gatgctcttc gtccagatca tcctgatcga caagaccggc ttccatccga 13680
gtacgtgctc gctcgatgcg atgtttcgtc tgggtggtcga atgggcaggt agccggatca 13740
agcgtatgca gccgccgat tgcacagcc atgatggata ctttctcggc aggagcaagg 13800
tgagatgaca ggagatcctg ccccggcact tcgccaata gcagccagtc cttcccgtc 13860
tcagtgacaa cgtcgagcac agctgcgcaa ggaacgcccg tcgtggccag ccacgatagc 13920
cgcgctgcct cgtcctgcag ttcattcagg gcaccggaca ggtcggctctt gacaaaaaga 13980
accgggcgcc cctgcgctga cagccggaac acggcggcat cagagcagcc gattgtctgt 14040
tgtgcccagt catagccgaa tagcctctcc acccaagcgg ccggagaacc tgcgtgcaat 14100
ccatcttgtt caatcatgcg aaacgatcca gatccggtgc agattatttg gattgagagt 14160
gaatatgaga ctctaattgg ataccgaggg gaatttatgg aacgtcagtg gagcattttt 14220
gacaagaaat atttgctagc tgatagtgac cttaggcgac ttttgaacgc gcaataatgg 14280
tttctgacgt atgtgcttag ctcatataac tccagaaacc cgcggctgag tggctccttc 14340

aacgttgccg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg 14400
tcataacgtg actcccttaa ttctccgctc atgatcagat tgctgtttcc cgccttcagt 14460
ttaaactatc agtgtttgac aggatatatt ggcgggtaaa cctaagagaa aagagcgttt 14520
attagaataa tcggatatatt aaaagggcgt gaaaagggtt atccgttcgt ccatttgtat 14580
gtgcatgcc aaccacaggt tccccagatc tggcgccggc cagcgagacg agcaagattg 14640
gccgccgccc gaaacgatcc gacagcgcg ccagcacagg tgcgcaggca aattgcacca 14700
acgcatacag cgccagcaga atgccatagt gggcggtgac gtcgttcgag tgaaccagat 14760
cgcgcaggag gcccggcagc accggcataa tcaggccgat gccgacagcg tcgagcgcg 14820
cagtgtcag aattacgatc aggggtatgt tgggtttcac gtctggcctc cggaccagcc 14880
tccgctggtc cgattgaacg cgcggttct ttatcactga taagttggtg gacatattat 14940
gtttatcagt gataaagtgt caagcatgac aaagttgcag ccgaatacag tgatccgtgc 15000
cgccctggac ctgttgaacg aggtcggcgt agacggctctg acgacacgca aactggcgga 15060
acggttgggg gttcagcagc cggcgcttta ctggcacttc aggaacaagc gggcgctgct 15120
cgacgcactg gccgaagcca tgctggcgga gaatcatac cattcggtgc cgagagccga 15180
cgacgactgg cgctcatttc tgatcgggaa tgcccgcagc ttcaggcagg cgctgctcgc 15240
ctaccgcgat ggcgcgcgca tccatgccgg cacgcgaccg ggcgcaccgc agatggaaac 15300
ggccgacgcg cagcttcgct tcctctgcga ggcgggtttt tcggccgggg acgccgtcaa 15360
tgcgctgatg acaatcagct acttcactgt tggggccgtg cttgaggagc aggcggcgga 15420
cagcgatgcc ggcgagcgcg gcggcaccgt tgaacaggct ccgctctcgc cgctgttgcg 15480
ggccgcgata gacgccttcg acgaagccgg tccggacgca gcgttcgagc agggactcgc 15540
ggtgattgtc gatggattgg cgaaaaggag gctcgttgct aggaacgttg aaggaccgag 15600

aaaggggtgac gattgatcag gaccgctgcc ggagcgcaac ccaactcacta cagcagagcc 15660
atgtagacaa catccctcc ccctttccac cgcgtcagac gcccgtagca gcccgctacg 15720
ggctttttca tgccctgccc tagcgtccaa gcctcaaggc cgcgctcggc ctctctggcg 15780
gccttctggc gctcttcgc ttctctgctc actgactcgc tgcgctcggc cgttcggctg 15840
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat 15900
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 15960
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc 16020
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 16080
agctccctcg tgcgctctcc tgttccgacc ctgccgtta ccggatacct gtccgccttt 16140
ctcccttcgg gaagcgtggc gcttttccgc tgcataacc tgcttcgggg tcattatagc 16200
gattttttcg gtatatccat cctttttcgc acgatataca ggattttgcc aaagggttcg 16260
tgtagacttt ccttggtgta tccaacggcg tcagccgggc aggataggtg aagtagggcc 16320
acccgcgagc ggggtgttct tcttactgt cccttattcg cacctggcgg tgctcaacgg 16380
gaatcctgct ctgcgaggct ggccggctac cgccggcgta acagatgagg gcaagcggat 16440
ggctgatgaa accaagccaa ccaggaaggg cagcccacct atcaaggtgt actgccttcc 16500
agacgaacga agagcgattg aggaaaaggc ggcggcggcc ggcattgagc tgtcggccta 16560
cctgctggcc gtcggccagg gctacaaaat cacgggcgtc gtggactatg agcacgtccg 16620
cgagctggcc cgcattcaatg gcgacctggc ccgcctgggc ggcttctga aactctggct 16680
caccgacgac ccgcgcacgg cgcggttcgg tgatgccacg atcctcgccc tgctggcgaa 16740
gatcgaagag aagcaggacg agcttggcaa ggtcatgatg ggcgtgggtcc gcccgagggc 16800
agagccatga ctttttttagc cgctaaaacg gccgggggggt gcgctgatt gccaaagcac 16860

tccccatgcg ctccatcaag aagagcgact tcgcgagact ggtgaagtac atcaccgacg 16920

agcaaggcaa gaccgagcgc ctttgcgacg ctca 16954

<210> 45

<211> 19491

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (18970)..(18970)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19178)..(19178)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19269)..(19269)

<223> n is a, c, g, or t

<400> 45

agcttggtac cgagctcgga tccactagta acggccgcca gtgtgctgga attcgccctt 60

gacggccagt gaattcgagc tcggtacccg gggatctttc gacactgaaa tacgtcgagc 120

ctgctccgct tggaagcggc gaggagcctc gtcctgtcac aactaccaac atggagtacg 180

ataagggccca gttccgccag ctcatthaaga gccagttcat gggcggtggc atgatggccg 240

tcatgcatct gtacttcaag tacaccaacg ctcttctgat ccagtcgatc atccgctgaa 300

ggcgctttcg aatctgggta agatccacgt cttcggggaag ccagcgactg gtgacctcca 360

gcgtcccttt aaggctgcca acagctttct cagccagggc cagcccaaga ccgacaaggc 420

ctccctccag aacgccgaga agaactggag ggggtgggtgc aaggaggagt aagctcctta 480

ttgaagtcgg aggacggagc ggtgtcaaga ggatattctt cgactctgta ttatagataa 540

gatgatgagg aattggagggt agcatagctt catttggatt tgctttccag gctgagactc 600

tagcttggag catagaggggt cctttggctt tcaatattct caagtatctc gagtttgaac 660

ttattccctg tgaacctttt attcaccaat gagcattgga atgaacatga atctgaggac 720

tgcaatcgcc atgaggtttt cgaaatacat ccggatgtcg aaggcttggg gcacctgcgt 780

tggttgaatt tagaacgtgg cactattgat catccgatag ctctgcaaag ggcgttgcac 840

aatgcaagtc aaacgttgct agcagttcca ggtggaatgt tatgatgagc attgtattaa 900

atcaggagat atagcatgat ctctagttag ctcaccacaa aagtcagacg gcgtaaccaa 960

aagtcacaca acacaagctg taaggatttc ggcacggcta cggaagacgg agaagccacc 1020

ttcagtggac tcgagtacca tttaattcta tttgtgtttg atcgagacct aatacagccc 1080

ctacaacgac catcaaagtc gtatagctac cagtgaggaa gtggactcaa atcgacttca 1140

gcaacatctc ctggataaac tttaagccta aactatacag aataagatag gtggagagct 1200

tataccgagc tcccaaattct gtccagatca tgggtgaccg gtgcctggat cttcctatag 1260

aatcatcctt attcgttgac ctagctgatt ctggagtgc ccagaggggtc atgacttgag 1320

cctaaaatcc gccgcctcca ccatttgtag aaaaatgtga cgaactcgtg agctctgtac 1380

agtgaccgggt gactctttct ggcatgcgga gagacggacg gacgcagaga gaagggctga 1440

gtaataagcc actggccaga cagctctggc ggctctgagg tgcagtggat gattattaat 1500

ccgggaccgg ccgcccctcc gccccgaagt ggaaaggctg gtgtgcccct cgttgaccaa 1560

gaatctattg catcatcgga gaatatggag cttcatcgaa tcaccggcag taagcgaagg 1620

agaatgtgaa gccaggggtg tatagccgtc ggcgaaatag catgccatta acctaggtac 1680
agaagtccaa ttgcttccga tctggtaaaa gattcacgag atagtacctt ctccgaagta 1740
ggtagagcga gtacccggcg cgtaagctcc ctaattggcc catccggcat ctgtagggcg 1800
tccaaatata gtgcctctcc tgctttgccc ggtgtatgaa accggaaagg ccgctcagga 1860
gctggccagc ggcgagacc gggaacacaa gctggcagtc gacctatccg gtgctctgca 1920
ctcgacctgc tgaggtccct cagtccctgg taggcagctt tgccccgtct gtccgcccgg 1980
tgtgtcggcg gggttgacaa ggtcgttgcg tcagtccaac atttgttgcc atattttcct 2040
gctctcccca ccagctgctc ttttcttttc tctttctttt cccatcttca gtatatcat 2100
cttcccatcc aagaaccttt atttccccta agtaagtact ttgctacatc catactccat 2160
ccttcccatc ccttattcct ttgaaccttt cagttcgagc tttcccactt catcgagct 2220
tgactaacag ctaccccgct tgagcagaca tcaccatgct gtcgaagctg cagtcaatca 2280
gcgtcaaggc ccgccgcgtt gaactagccc gcgacatcac gcggcccaaa gtctgcctgc 2340
atgctcagcg gtgctcgta gttcggctgc gagtggcagc accacagaca gaggaggcgc 2400
tggaaccgt gcaggctgcc ggcgggggcg atgagcacag cgccgatgta gcactccagc 2460
agcttgaccg ggctatcgca gagcgctgtg cccggcgcaa acgggagcag ctgtcatacc 2520
aggctgccgc cattgcagca tcaattggcg tgtcaggcat tgccatcttc gccacctacc 2580
tgagatttgc catgcacatg accgtgggcg gcgcagtgcc atggggtgaa gtggctggca 2640
ctctcctctt ggtggttggt ggcgcgctcg gcatggagat gtatgccgc tatgcacaca 2700
aagccatctg gcatgagtcg cctctgggct ggctgctgca caagagccac cacacacctc 2760
gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc 2820
tcctgtgtac ctttggcttc tggctgccca acgtcctggg ggcggcctgc tttggagcgg 2880

ggctgggcat cagctatac ggcattggcat atatgtttgt acacgatggc ctggtgcaca 2940
ggcgctttcc caccggggccc atcgctggcc tgccctacat gaagcgctg acagtggccc 3000
accagctaca ccacagcggc aagtacgggtg gcgcgccttg gggatatgtc ttgggtccac 3060
aggagctgca gcacattcca ggtgcggcgg aggaggtgga gcgactggtc ctggaactgg 3120
actggtccaa gcggtagggt gcggaaccag gcacgtggt ttcacacctc atgcctgtga 3180
taagggtggtg ctagagcgat gcgtgtgaga cgggtatgtc acggtcgact ggtctgatgg 3240
ccaatggcat cggccatgtc tggatcatcac gggctggttg cctgggtgaa ggtgatgcac 3300
atcatcatgt gcggttgag gggctggcac agtgtgggct gaactggagc agttgtccag 3360
gctggcggtt aatcagtgag ggtttgtgat tggcggttgt gaagcaatga ctccgcccac 3420
attctatttg tgggagctga gatgatggca tgcttgggat gtgcatggat catggtagtg 3480
cagcaaacta tattcaccta gggctgttg taggatcagg tgaggccttg cacattgcat 3540
gatgtactcg tcatggtgtg ttggtgagag gatggatgtg gatggatgtg tattctcaga 3600
cgtagacctt gactggaggc ttgatcgaga gagtgggccc tattctttga gaggggaggc 3660
tcgtgccaga aatggtgagt ggatgactgt gacgtgtac attgcaggca ggtgagatgc 3720
actgtctcga ttgtaaaata cattcagatg caagcttggc gtaatcatgg tcatagctgt 3780
ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcataa 3840
agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaatgctg ttgcgctcac 3900
tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgctg 3960
cggggagagg cggtttgcgt attgggcca agacaaaagg gcgacattca accgattgag 4020
ggaggggaagg taaatattga cggaattat tcattaaagg tgaattatca ccgtcaccga 4080
cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4140

ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat 4200
caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg gtcataagccc 4260
ccttatttagc gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg 4320
aacgcctcc ctcagagccg ccaccctcag aaccgccacc ctcagagcca ccaccctcag 4380
agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc ccgatctagt 4440
aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat ttgttttct 4500
atcgcgatatt aatgtataa ttgcgggact ctaatcataa aaaccatct cataaataac 4560
gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4620
atcatcgcaa gaccggcaac aggattcaat cttaagaaac tttattgcca aatgtttgaa 4680
cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 4740
atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt gatgtggacg 4800
ccggggccga tcatattgtc gctcaggatc gtggcgttgt gcttgtcggc cgttgctgtc 4860
gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc gaagaactcc 4920
agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac gattccgaag 4980
cccaaccttt catagaaggc ggcgggtggaa tcgaaatctc gtgatggcag gttgggcgtc 5040
gcttggtcgg tcatttcgaa ccccagagtc ccgctcagaa gaactcgtca agaaggcgat 5100
agaaggcgat gcgctgcgaa tcggggagcgg cgataccgta aagcacgagg aagcggtcag 5160
cccattcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg tcctgatagc 5220
ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca tttccacca 5280
tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc 5340
gcgccttgag cctggcgaaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat 5400

catcctgatac gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 5460

cttgggtggc gaatgggcag gtagccggat caagcgtatg cagccgcgc attgcatcag 5520

ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca 5580

cttcgccccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc acagctgcgc 5640

aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctgcctctgc agttcattca 5700

gggcaccgga caggtcggc ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga 5760

acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagcg aatagcctct 5820

ccacccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgatc 5880

cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 5940

gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg 6000

accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agtcattaa 6060

actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca gttccaaacg 6120

taaaacggct tgtccgcgt catcggcggg ggtcataacg tgactccctt aattctccgc 6180

tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg acaggatata 6240

ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcgatat ttaaaagggc 6300

gtgaaaagggt ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga 6360

tctggcgccg gccagcgaga cgagcaagat tggccgcgc ccgaaacgat ccgacagcgc 6420

gccagcaca ggtgcgcagg caaattgcac caacgatac agcgccagca gaatgccata 6480

gtgggcggtg acgtcgttcg agtgaaccag atcgcgagg aggcccgga gcaccggcat 6540

aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga tcaggggtat 6600

gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggatt 6660

ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg 6720
acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggctcggc 6780
gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca gccggcgctt 6840
tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc catgctggcg 6900
gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt tctgatcggg 6960
aatgcccga gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg catccatgcc 7020
ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg ctccctctgc 7080
gaggcgggtt tttcggccgg ggacgccgtc aatgcgctga tgacaatcag ctacttcaact 7140
gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc 7200
gttgaacagg ctccgctctc gccgctgttg cgggcgcga tagacgcctt cgacgaagcc 7260
ggtccggacg cagcgttcga gcagggaactc gcggtgattg tcgatggatt ggcgaaaagg 7320
aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc aggaccgctg 7380
ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct ccccctttcc 7440
accgcgtcag acgcccgtag cagcccgtc cgggcttttt catgccctgc cctagcgtcc 7500
aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttctctgc 7560
tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 7620
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 7680
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 7740
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcgacag 7800
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 7860
ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgcttttcc 7920

gctgcataac cctgcttcgg ggtcattata gcgatTTTTT cggatatatcc atcctTTTTT 7980

gcacgatata caggatTTTT ccaaagggtt cgtgtagact ttccttggtg tatccaacgg 8040

cgtcagccgg gcaggatagg tgaagtaggc ccacccgga gcggtgttc cttcttctact 8100

gtcccttatt cgcacctggc ggtgctcaac gggaaatcctg ctctgcgagg ctggccggct 8160

accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag 8220

ggcagccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag 8280

gcggcggcgg ccggcatgag cctgtcgcc tacctgctgg ccgtcggcca gggctacaaa 8340

atcacggcg tcgtggacta tgagcacgtc cgcgagctgg ccgcatcaa tggcgacctg 8400

ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgac ggcgcggttc 8460

ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc 8520

aaggtcatga tgggcgtggt ccgcccagg gcagagccat gactTTTTTA gccgctaaaa 8580

cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga 8640

cttcgcggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcctttgcga 8700

cgctcaccgg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc cctgcaaacg 8760

cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcg ccgccggcgt tgtggatacc 8820

tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact tgaggggccg 8880

actcaccgg cgcggcgttg acagatgagg ggcaggctcg atttcggccg gcgacgtgga 8940

gctggccagc ctgcgaaatc ggcgaaaac cctgatttta cgcgagtttc ccacagatga 9000

tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg 9060

acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgaggggcgc 9120

acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc aagggtttcc 9180

gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa 9240
accttgtttt taaccagggc tgcgcctgt gcgcgtgacc gcgcacgccg aaggggggtg 9300
cccccccttc tcgaaccctc ccggcccgt aacgcgggccc tcccatcccc ccaggggctg 9360
cgccccctcg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt 9420
gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc cggaagcatt 9480
gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc 9540
ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga cttcatggcg 9600
gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc cgtgctcgtg 9660
ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt ataccgaggt 9720
atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat ttaaaaagct 9780
accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat attgacaata 9840
ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg 9900
caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaaagca taaaaacttg 9960
catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt ctatcataat 10020
tggttaatga ctccaactta ttgatagtgt tttatgttca gataatgccc gatgactttg 10080
tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccaggtgc 10140
tgctcagat tcaggttatg ccgtcaatt cgctgcgtat atcgcttgct gattacgtgc 10200
agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca tatcaccag 10260
tcaaaggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg ttcaccgaat 10320
acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca gcgtggcg 10380
gatttagccc cgacatagcc cactgttcg tccatttccg cgagacgat gacgtcactg 10440

cccggtgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga cgtaaaatcg 10500
tggtgaggcc aacgcccata atgcgggctg ttgcccgga tccaacgcca ttcattggcca 10560
tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac tgcagttgcc 10620
atgttttacg gcagtgcag cagagatagc gctgatgtcc ggcggtgctt ttgccgttac 10680
gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10740
cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cataattgtg 10800
gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac aactttgaaa 10860
aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct 10920
tggtataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 10980
aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11040
aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag aaaatgaaaa 11100
cctatattta aaaatgacgg acagccggt taaagggacc acctatgatg tggaacggga 11160
aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtc tgcactttga 11220
acggcatgat ggctggagca atctgctcat gaggaggcc gatggcgctc ttgctcgga 11280
agagtatgaa gatgaacaaa gccctgaaa gattatcgag ctgtatgcgg agtgcacag 11340
gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt 11400
agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga 11460
agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaagacc 11520
cgaagaggaa cttgtctttt ccacggcga cctgggagac agcaacatct ttgtgaaaga 11580
tggaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtggtatga 11640
cattgccttc tgcgtccggt cgatcaggga ggatcggg gaagaacagt atgtcgagct 11700

atatttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atatttttact 11760
ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 11820
ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc aagtatttgg 11880
gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac gagaaggacg 11940
gccagacggg ctacggggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12000
caccaggcgg gtcaaatcag gaataagggc acattgcccc ggcgtgagtc ggggcaatcc 12060
cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12120
acgcgggggt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc atgcgtgcgc 12180
cccgcgaaac cttccagtcc gtcggctcga tggccagca agctacggcc aagatcgagc 12240
gcgacagcgt gcaactggct cccctgccc tgcccgcgc atcgccgcgc gtggagcgtt 12300
cgcgctcgtc cgaacaggag gcggcagggt tggcgaagtc gatgaccatc gacacgcgag 12360
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg 12420
aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt 12480
ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaaac gacacggccc 12540
gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg caaaacaagg 12600
tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag ctgcgggccg 12660
acgatgacga actggtgtgg cagcagggtg tggagtacgc gaagcgcacc cctatcggcg 12720
agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggctg atcaatggcc 12780
gggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg atgggcttca 12840
cgtccgaccg cggtgggcac ctggaatcgg tgtcgtgct gcaccgcttc cgcgtcctgg 12900
accgtggcaa gaaaacgtcc cggtgccagg tcctgatcga cgaggaaatc gtcgtgctgt 12960

ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg 13020
cccgaaggat gttcgactat ttcagctcgc accgggagcc gtacccgctc aagctggaaa 13080
ccttccgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc gagcaggtcg 13140
gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg gtcaatgatg 13200
acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg gggttcagcag 13260
ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact tgcttcgctc 13320
agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag gattaaaatt 13380
gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc aggatttccg 13440
cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg tttacgagca 13500
cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg tggcattcgg 13560
cgctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg acggcccca 13620
ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc gaggccgagg 13680
ggtcgccggt atgctgctgc gggcgcttgc ggcgggttta ttgctcgtga tgatcgctcg 13740
acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac ttaatatctc 13800
gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccgggcgggg tcgcggcgac 13860
ggtaggcgct gtgcagccgc tgatggtcgt gttcatctct gccgctctgc taggtagccc 13920
gatacgattg atggcgggcc tgggggctat ttgcggaact gcgggcgtgg cgctgttggg 13980
gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg cgggggcggg 14040
ttccatggcg ttcggaaccg tgctgacctg caagtggcaa cctcccgtgc ctctgctcac 14100
ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag ctttagtggt 14160
tgatccgcca atcccgatgc ctacaggaac caatgttctc ggcctggcgt ggctcggcct 14220

gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac tcgaacctac 14280
agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc cggggatgca 14340
tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag caatggatag 14400
gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc ttcctcagcg 14460
gctttatcca gcgatttcct attatgtcgg catagtcttc aagatcgaca gcctgtcacg 14520
gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg agatgatatt 14580
tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 14640
atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14700
gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 14760
ctgtatcgag tgggtgatttt gtgccgagct gccggtcggg gagctgttgg ctggctggtg 14820
gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac acattgcgga 14880
cgtttttaat gtactggggg ggtttttctt ttcaccagtg agacgggcaa cagctgattg 14940
cccttcaccg cctggccctg agagagttgc agcaagcggc ccacgctggg ttgccccagc 15000
aggcgaaaat cctgtttgat ggtgggttcg aaatcggcaa aatcccttat aaatcaaaag 15060
aatagcccga gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga 15120
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 15180
aaccatcacc caaatcaagt tttttggggg cgaggtgccg taaagcacta aatcggaacc 15240
ctaaaggag cccccgattt agagcttgac ggggaaagcc ggcgaaacgtg gcgagaaagg 15300
aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg ggaagggcga 15360
tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 15420
ttaagttggg taacgccagg gttttccag tcacgacgtt gtaaaacgac ggccagtga 15480

ttcgagctcg gtacccgggg atctttcgac actgaaatac gtcgagcctg ctccgcttgg 15540
aagcggcgag gagcctcgtc ctgtcacaac taccaacatg gagtacgata agggccagtt 15600
ccgccagctc attaagagcc agttcatggg cgttggcatg atggccgtca tgcattctgta 15660
cttcaagtac accaacgctc ttctgatcca gtcgatcctc cgctgaaggc gctttcgaat 15720
ctggttaaga tccacgtctt cgggaagcca gcgactggtg acctccagcg tccctttaag 15780
gctgccaaca gctttctcag ccagggccag cccaagaccg acaaggcctc cctccagaac 15840
gccgagaaga actggagggg tgggtgtcaag gaggagtaag ctcccttattg aagtcggagg 15900
acggagcggg gtcaagagga tattcttcga ctctgtatta tagataagat gatgaggaat 15960
tggaggtagc atagcttcat ttggatttgc tttccaggct gagactctag cttggagcat 16020
agagggctct ttggctttca atattctcaa gtatctcgag tttgaactta ttccctgtga 16080
accttttatt caccaatgag cattggaatg aacatgaatc tgaggactgc aatcgccatg 16140
agggttttga aatacatccg gatgtcgaag gcttggggca cctgcgcttg ttgaatttag 16200
aacgtggcac tattgatcat ccgatagctc tgcaaagggc gttgcacaat gcaagtcaaa 16260
cgttgctagc agttccaggg ggaatgttat gatgagcatt gtattaaatc aggagatata 16320
gcatgatctc tagttagctc accacaaaag tcagacggcg taaccaaag tcacacaaca 16380
caagctgtaa ggatttcggc acggctacgg aagacggaga agccaccttc agtggactcg 16440
agtaccattt aattctattt gtgtttgatc gagacctaat acagccccta caacgaccat 16500
caaagtcgta tagctaccag tgaggaagtg gactcaaatc gacttcagca acatctcctg 16560
gataaacttt aagcctaaac tatacagaat aagataggtg gagagcttat accgagctcc 16620
caaactctgc cagatcatgg ttgaccgggtg cctggatctt cctatagaat catccttatt 16680
cgttgacctc gctgattctg gactgacca gaggggtcatg acttgagcct aaaatccgcc 16740

gcctccacca tttgtagaaa aatgtgacga actcgtgagc tctgtacagt gaccggtgac 16800
tctttctggc atgcggagag acggacggac gcagagagaa gggctgagta ataagccact 16860
ggccagacag ctctggcggc tctgaggtgc agtggatgat tattaatccg ggaccggccg 16920
ccccccgccc ccgaagtgga aaggctgggtg tgccccctcgt tgaccaagaa tctattgcat 16980
catcggagaa tatggagctt catcgaatca ccggcagtaa gcgaaggaga atgtgaagcc 17040
aggggtgtat agccgtcggc gaaatagcat gccattaacc taggtacaga agtccaattg 17100
cttcogatct ggtaaaagat tcacgagata gtaccttctc cgaagtaggt agagcgagta 17160
cccggcgcgt aagctcccta attggcccat ccggcatctg tagggcgctc aaatatcgtg 17220
cctctcctgc tttgcccggt gtatgaaacc ggaaaggccg ctcaggagct ggccagcggc 17280
gcagaccggg aacacaagct ggcagtcgac ccatccgggtg ctctgcactc gacctgctga 17340
ggtccctcag tccctggtag gcagctttgc cccgtctgtc cgcccgggtgt gtcggcgggg 17400
ttgacaaggt cgttgogtca gtccaacatt tgttgccata ttttctgct ctccccacca 17460
gctgctcttt tcttttctct ttcttttccc atcttcagta tattcatctt cccatccaag 17520
aacctttatt tcccctaagt aagtactttg ctacatccat actccatcct tcccatccct 17580
tattcctttg aacctttcag ttcgagcttt cccacttcat cgcagcttga ctaacagcta 17640
ccccgcttga gcagacatca ccatgcctga actcaccgcg acgtctgtcg agaagtttct 17700
gatcgaaaag ttcgacagcg tctccgacct gatgcagctc tcggagggcg aagaatctcg 17760
tgctttcagc ttcgatgtag gagggcgtgg atatgtcctg cgggtaaata gctgcgccga 17820
tggtttctac aaagatcggt atgtttatcg gcactttgca tcggccgcgc tcccgatcc 17880
ggaagtgctt gacattgggg aattcagcga gagcctgacc tattgcatct cccgccgtgc 17940
acaggggtgc acgttgcaag acctgcctga aaccgaactg cccgctgttc tgcagccggt 18000

cgcgagggcc atggatgcga tcgctgcggc cgatcttagc cagacgagcg ggttcggccc 18060
attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat gcgcgattgc 18120
tgatcccat gtgtatcact ggcaaactgt gatggacgac accgtcagtg cgtccgtcgc 18180
gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc ggcacctcgt 18240
gcacgcggat ttcggctcca acaatgtcct gacggacaat ggccgcataa cagcggtcatt 18300
tgactggagc gaggcgatgt tcggggattc ccaatacgag gtcgccaaca tcttcttctg 18360
gaggccgtgg ttggcttgta tggagcagca gacgcgctac ttcgagcggg ggcattccgga 18420
gcttgcagga tcgccgcggc tccggggcgta tatgctccgc attggtcttg accaactcta 18480
tcagagcttg gttgacggca atttcgatga tgcagcttgg gcgcagggtc gatgcgacgc 18540
aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca gaagcgcggc 18600
cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac gcccagcac 18660
tcgtccgagg gcaaaggaat agagtagatg ccgaccgcgg gatcgatcca cttaacgtta 18720
ctgaaatcat caaacagctt gacgaatctg gatataagat cgttggtgtc gatgtcagct 18780
ccggagtga gacaaatggg gttcaggatc tcgataagat acgttcattt gtccaagcag 18840
caaagagtgc cttctagtga tttaatagct ccatgtcaac aagaataaaa cgcgttttcg 18900
ggtttacctc ttccagatac agctcatctg caatgcatta atgcattgac tgcaacctag 18960
taacgccttn caggctccgg cgaagagaag aatagcttag cagagctatt ttcattttcg 19020
ggagacgaga tcaagcagat caacggctgt caagagacct acgagactga ggaatccgct 19080
cttggtcca cgcgactata tatttgtctc taattgtact ttgacatgct cctcttcttt 19140
actctgatag cttgactatg aaaattccgt caccagcncc tgggttcgca aagataattg 19200
catgtttctt ccttgaactc tcaagcctac aggacacaca ttcatcgtag gtataaacct 19260

cgaaatcant tcctactaag atggtataca atagtaacca tgcátgggtg cctagtgaat 19320
gctccgtaac acccaatacg ccggccgaaa cttttttaca actctcctat gagtcgttta 19380
cccagaatgc acaggtacac ttgttttagag gtaatccttc tttctagcta gaagtcctcg 19440
tgtactgtgt aagcgcccac tccacatctc cactcgacct gcaggcatgc a 19491

<210> 46
<211> 21300
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 46
gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60
cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120
cagttcatgg gcgttggcat gatggccgct atgcatctgt acttcaagta caccaacgct 180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctgggtaag atccacgtct 240

tcgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca 300

gccagggcca gccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360

gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg 420

atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

tttgatttg ctttccaggc tgagactcta gcttgaggca tagagggtcc tttggcttcc 540

aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600

gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660

ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720

tccgatagct ctgcaaaggg cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780

tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaacttgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcatccttat tcgttgacct agctgattct 1200

ggagtgacct agagggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgagggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcatcggaga atatggagct 1500

tcatacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560
cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgatc tggtaaaaga 1620
ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680
aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccgg 1740
tgtatgaaac cggaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800
tggcagtcga cccatccggg gctctgcact cgacctgctg aggtccctca gtccctggta 1860
ggcagctttg ccccgctctg ccgcccgggtg tgtcggcggg gttgacaagg tcgttgcgtc 1920
agtccaacat ttgttgccat attttctgc tctcccacc agctgctctt ttcttttctc 1980
tttcttttcc catcttcagt atattcatct tcccatcaa gaacctttat ttcccctaag 2040
taagtacttt gctacatcca tactccatcc tcccatccc ttattccttt gaacctttca 2100
gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctcggagggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcggtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340
tatgtttatc ggcactttgc atcgcccgcg ctcccgattc cggaagtgtc tgacattggg 2400
gaattcagcg agagcctgac ctattgcac tcccgccgtg cacaggggtg cacgttgcaa 2460
gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580
ggtaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg cccgaagtc cggcacctcg tgcacgcgga tttcggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820
ttcgggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgcagg atcgccgcgg 2940
ctccggggcgt atatgctccg cattgggtctt gaccaactct atcagagctt ggttgacggc 3000
aatttcgatg atgcagcttg ggcgcagggg cgatgcgacg caatcgctccg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgccccg agaagcgcgg ccgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cggccagca ctcgctccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ctttctagt 3360
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaacct a gtaacgcctt ncagggtccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600
atatttgtct ctaattgtac tttgacatgc tcctcttctt tactctgata gcttgactat 3660
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780
gatgggtatac aatagtaacc atgcatgggt gcctagttaa tgctccgtaa cacccaatac 3840
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggatca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccaactcgacc tgcaggcatg caagcttgaa ttcgagctcg gtacccgggg 4020

atcttttcgac actgaaatac gtcgagcctg ctccgcttgg aagcggcgag gagcctcgtc 4080
ctgtcacaac taccaacatg gagtacgata agggccagtt ccgccagctc attaagagcc 4140
agttcatggg cgttggcatg atggccgtca tgcattctgta cttcaagtac accaacgctc 4200
ttctgatcca gtcgatcatc cgctgaaggc gctttcgaat ctggttaaga tccacgtctt 4260
cgggaagcca gcgactggtg acctccagcg tccctttaag gctgccaaca gctttctcag 4320
ccagggccag cccaagaccg acaaggcctc cctccagaac gccgagaaga actggagggg 4380
tggtgtcaag gaggagtaag ctcccttattg aagtcggagg acggagcggg gtcaagagga 4440
tattcttcga ctctgtatta tagataagat gatgaggaat tggaggtagc atagcttcat 4500
ttggatttgc tttccaggct gagactctag cttggagcat agagggctct ttggctttca 4560
atattctcaa gtatctcgag tttgaactta ttccctgtga accttttatt caccaatgag 4620
cattggaatg aacatgaatc tgaggactgc aatcgccatg aggttttcga aatacatccg 4680
gatgtcgaag gcttggggca cctgcgttgg ttgaatttag aacgtggcac tattgatcat 4740
ccgatagctc tgcaaagggc gttgcacaat gcaagtcaaa cgttgctagc agttccagg 4800
ggaatgttat gatgagcatt gtattaaatc aggagatata gcatgatctc tagttagctc 4860
accacaaaag tcagacggcg taaccaaag tcacacaaca caagctgtaa ggatttcggc 4920
acggctacgg aagacggaga agccaccttc agtggactcg agtaccattt aattctattt 4980
gtgtttgatc gagaccta acagccccta caacgaccat caaagtcgta tagctaccag 5040
tgaggaagtg gactcaaatc gacttcagca acatctcctg gataaacttt aagcctaaac 5100
tatacagaat aagataggtg gagagcttat accgagctcc caaatctgtc cagatcatgg 5160
ttgaccggtg cctggatctt cctatagaat catccttatt cgttgacctt gctgattctg 5220
gagtgacca gaggggtcatg acttgagcct aaaatccgcc gcctccacca tttgtagaaa 5280

aatgtgacga actcgtgagc tctgtacagt gaccggtgac tctttctggc atgcggagag 5340
acggacggac gcagagagaa gggctgagta ataagccact ggccagacag ctctggcggc 5400
tctgaggtgc agtggatgat tattaatccg ggaccggccg cccctccgcc ccgaagtgga 5460
aaggctggtg tgcccctcgt tgaccaagaa tctattgcat catcggagaa tatggagctt 5520
catcgaatca ccggcagtaa gcgaaggaga atgtgaagcc aggggtgtat agccgtcggc 5580
gaaatagcat gccattaacc taggtacaga agtccaattg cttccgatct ggtaaaagat 5640
tcacgagata gtaccttctc cgaagtaggt agagcgagta cccggcgcggt aagctcccta 5700
attggcccat ccggcatctg tagggcgctc aaatatcgtg cctctcctgc tttgcccggc 5760
gtatgaaacc ggaaaggccg ctcaggagct ggccagcggc gcagaccggg aacacaagct 5820
ggcagtcgac ccatccgggtg ctctgcactc gacctgctga ggtccctcag tccctggtag 5880
gcagctttgc cccgtctgtc cgcccgggtg gtcggcgggg ttgacaaggt cgttgcgta 5940
gtccaacatt tgttgccata ttttctgct cccccacca gctgctctt tcttttctct 6000
ttcttttccc atcttcagta tattcatctt cccatccaag aacctttatt tcccctaagt 6060
aagtactttg ctacatccat actccatcct toccatccct tattcctttg aacctttcag 6120
ttcgagcttt cccacttcat cgcagcttga ctaacagcta ccccgcttga gcagacatca 6180
ccatgtcaat actcacttat ctggaatttc atctctacta tacactacct gtccttgccg 6240
cattgtgttg gctgctaaag ccgtttcact cacagcaaga caatctcaag tataaatTTT 6300
taatgttgat ggccgcctct accgcatcga tttgggacaa ttatatcgtt tatcatcgcg 6360
cttggtggta ctgtcctact tgtgttggtg ctgtcattgg ctatgtacct ctagaagaat 6420
acatgttctt tatcatcatg actttaatga ctgtcgcgtt ctcaaacttt gttatgcgtt 6480
ggcaattgca tactttcttt attagacca acacttcttg gaagcaaaca ctattagtac 6540

gccttgtgcc tgtttcagct ttattggcaa tcacttatca tgcttggcac ttgacactgc 6600
caaataaacc ttcatTTTTat ggttcacgca tcctttggta tgcttgtcct gtgttggcta 6660
ttctttggct ggggtgctggc gaatatatct tgcgtcgacc tgtggctgtc cttttgtcta 6720
ttgttatccc tagtgatac ctatgttggg ctgatatcgt cgctattagt gctggcacat 6780
ggcatatttc tcttagaaca agcactggca aaatggtagt acccgattta cctgtagaag 6840
aatgcctgtt ttttactttg atcaacacag tcttggtttt tgctacctgt gctatagacc 6900
gcgctcaggc catcctccat gtgagcgcg gtaatacgac tcactatagg gcgaattgga 6960
gctccaccgc ggtggcggcc gctctagaac tagtggatcc cccgggctgc aggaattcgg 7020
cacgagctac atttcacaag cccgtgagcg gtgcaagcg tctgccccac atcggccac 7080
ctcctcatct ccacgggtca tttgctgcta ccacgatgct gtcgaagctg cagtcaatca 7140
gcgtcaaggc ccgcccgtt gaactagccc gcgacatcac gcggcccaaa gtctgcctgc 7200
atgctcagcg gtgctcgta gttcggctgc gagtggcagc accacagaca gaggaggcg 7260
tggaaccgt gcaggctgcc ggcgcggcg atgagcacag cgccgatgta gcactccagc 7320
agcttgaccg ggctatcgca gagcgtcgtg cccggcgcaa acgggagcag ctgtcatacc 7380
aggetgccgc cattgcagca tcaattggcg tgctaggcat tgccatcttc gccacctacc 7440
tgagatttgc catgcacatg accgtggcg gcgcagtgcc atggggtgaa gtggctggca 7500
ctctctctt ggtggttggg ggcgcgctcg gcatggagat gtatgcccg tatgcacaca 7560
aagccatctg gcatgagtcg cctctgggct ggctgctgca caagagccac cacacacctc 7620
gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc 7680
tcctgtgtac ctttggcttc tggctgcca acgtcctggg ggcggcctgc tttggagcgg 7740
ggctgggcat cacgctatac ggcattggcat atatgtttgt acacgatggc ctggtgcaca 7800

ggcgctttcc caccgggccc atcgctggcc tgccctacat gaagcgctg acagtggccc 7860
accagctaca ccacageggc aagtacgggtg gcgcgccttg gggatatgtc ttgggtccac 7920
aggagctgca gcacattcca ggtgcggcgg aggaggtgga gcgactggtc ctggaactgg 7980
actgggtccaa gcgggctcag gccatcctcc atctgtacaa atcatctgtt caaaatcaaa 8040
accctaaaca agccatttcc cttttccagc atgtcaaaga gctagcatgg gccttctgtc 8100
ttcctgacca aatgctcaac aatgaattgt ttgatgatct tactatcagc tgggatattt 8160
tacgtaaagc ctcaaagtca ttctatactg catctgccgt tttccaagt tatgtacgtc 8220
aagacttggg tgttctctat gctttctgca gagctaccga tgacctgtgc gatgatgaat 8280
ccaaatctgt tcaagaaaga agagaccaat tagatcttac tcgacaattt gttcgtgatc 8340
tcttttagcca aaagaccagt gcgcctattg tgattgattg ggaattgtat caaaaccaac 8400
ttcctgcttc ttgtatatca gccttttagag cctttactcg ccttcgccat gtccttgaag 8460
tagaccctgt agaagaacta ttagatgggtt acaaattggga tcttgagcgt cgtcctatcc 8520
ttgatgaaca agacttggag gcatactctg cttgtgtggc cagtagtgtg ggtgaaatgt 8580
gcacacgtgt gattcttgct caagaccaa aggaaatga tgcttgata attgaccgtg 8640
cacgtgagat ggggctgggt ctacaatacg ttaacattgc tcgagacatt gtgactgata 8700
gcgagactct gggtcgatgt tatctgcctc aacaatggct tagaaaagaa gaaacagaac 8760
aaatacagca aggcaacgcc cgtagcctag gtgatcaaag actgttgggc ttgtctctga 8820
agctttagg aaaggcagac gctatcatgg tgagagctaa gaagggcatt gacaagttgc 8880
cggcaaactg tcaaggcggg gtacgagctg cttgccaaagt atatgctgca attggatctg 8940
tactcaagca gcagaagaca acatataccta caagagctca tctaaaagga agcgaacgtg 9000
ccaagattgc tctgttgagt gtatacaacc tctatcaatc tgaagacaag cctgtggctc 9060

tccgtcaagc tagaaagatt aagagttttt ttgttgatta gtgaattttt gttttattta 9120
tgtctgatag ttcaataaag agacaacaca tacaatataa aatcattgtc tttaaatgtt 9180
aatttagtag agtgtaaagc ctgcattttt tttgtacgca taaacaatga gttcaccccg 9240
cttctggttt ttaaataatt atgtcaaact agggaaaatt cttttttttc tcttcgttct 9300
ttttttggct tggttgaggag tcacaggctt gtcttcagat tgatagaggt tgtatacact 9360
caacagagca atcttggcac gttcgcttcc ttttagatga gctctttag gatatgttgt 9420
cttctgctgc ttgagtacag atccaattgc agcatatact tggcaagcag ctcgtacacc 9480
gccttgacag tttgccggca acttgtaaat gcccttctta gctctcacca tgatagcgtc 9540
tgcctttcct acaagcttgg cgtaatcatg gtcatactgt tttcctgtgt gaaattgtta 9600
tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag cctgggggtgc 9660
ctaagtgtg agctaactca cattaattgc gttgcgctca ctgcccgctt tccagtcggg 9720
aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag gcgggtttgcg 9780
tattgggcca aagacaaaag ggcgacattc aaccgattga gggaggggaag gtaaattattg 9840
acggaaatta ttcattaaag gtgaattatc accgtcaccg acttgagcca tttgggaatt 9900
agagccagca aaatcaccag tagcaccatt accattagca aggccggaaa cgtcaccaat 9960
gaaaccatcg atagcagcac cgtaatcagt agcgacagaa tcaagtttgc ctttagcgtc 10020
agactgtagc gcgttttcat cggcattttc ggtcatagcc cccttattag cgtttgccat 10080
cttttcataa tcaaaatcac cggaaccaga gccaccaccg gaaccgcctc cctcagagcc 10140
gccaccctca gaaccgccac cctcagagcc accaccctca gagccgccac cagaaccacc 10200
accagagccg ccgccagcat tgacaggagg cccgatctag taacatagat gacaccgcgc 10260
gcgataattt atcctagttt gcgcgctata ttttgttttc tatcgcgat taaatgtata 10320

attgcgggac tctaatacata aaaacccatc tcataaataa cgatcatgcat tacatgttaa 10380
ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca agaccggcaa 10440
caggattcaa tcttaagaaa ctttattgcc aaatgtttga acgatcgggg atcatccggg 10500
tctgtggcgg gaactccacg aaaatatccg aacgcagcaa gatatcgcg tgcatctcgg 10560
tcttgccctgg gcagtcgccg ccgacgccgt tgatgtggac gccgggccc atcatattgt 10620
cgctcaggat cgtggcggtg tgcttgctcg ccgttgctgt cgtaatgata tcggcacctt 10680
cgaccgcctg ttccgcagag atcccgtggg cgaagaactc cagcatgaga tccccgcgt 10740
ggaggatcat ccagccggcg tcccggaana cgattccgaa gcccaacctt tcatagaagg 10800
cggcgggtga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg gtcatttcga 10860
acccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga tgcgctgcga 10920
atcgggagcg gcgataccgt aaagcacgag gaagcgggtca gccattcgc cgccaagctc 10980
ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggtcgccca caccagccg 11040
gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg gcaagcaggc 11100
atcgccatgg gtcacgacga gatcatgcc gtcgggcatg cgcgcttga gcctggcgaa 11160
cagttcggct ggcgcgagcc cctgatgctc ttgctccaga tcctcctgat cgacaagacc 11220
ggcttccatc cgagtacgtg ctgctcgat gcgatgtttc gcttggtggg cgaatgggca 11280
ggtagccgga tcaagcgtat gcagccgccg cattgcatca gccatgatgg atactttctc 11340
ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgcccc atagcagcca 11400
gtcccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc ccgtcgtggc 11460
cagccacgat agccgcgctg cctcgtcctg cagttcattc agggcaccgg acaggtcggg 11520
cttgacaaaa agaaccgggc gccctgcgc tgacagccgg aacacggcgg catcagagca 11580

gccgattgtc tgttgtgccc agtcatagcc gaatagcctc tccacccaag cggccggaga 11640
acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg tgcagattat 11700
ttggattgag agtgaatatg agactctaata tggataccga ggggaattta tggaacgtca 11760
gtggagcatt ttgacaaga aatatttgct agctgatagt gaccttaggc gacttttgaa 11820
cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa acccgcggt 11880
gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaaacggc ttgtcccgcg 11940
tcatcggcgg gggtcataac gtgactccct taattctccg ctcgatgaca gattgtcggt 12000
tccgccttc agtttaaact atcagtgttt gacaggatat attggcgggt aaacctaga 12060
gaaaagagcg ttattagaa taatcggata tttaaaggc cgtgaaaagg ttatccgtt 12120
cgtccatttg tatgtgcatg ccaaccacag ggtccccag atctggcgcc ggccagcgag 12180
acgagcaaga ttggccgcg cccgaaacga tccgacagcg cggccagcac aggtgcgcgag 12240
gcaaattgca ccaacgcata cagcgccagc agaatgccat agtgggcgggt gacgtcggtc 12300
gagtgaacca gatcgcgag gagggccggc agcaccggca taatcaggcc gatgccgaca 12360
gcgtcgagcg cgacagtgtc cagaattacg atcaggggta tgttgggttt cacgtctggc 12420
ctccggacca gcctccgtg gtccgattga acgcgcggt tctttatcac tgataagttg 12480
gtggacatat tatgtttatc agtgataaag tgtcaagcat gacaaagttg cagccgaata 12540
cagtgatccg tgccgccctg gacctgttga acgaggtcgg cgtagacgggt ctgacgacac 12600
gcaaactggc ggaacgggtg ggggttcagc agccggcgct ttactggcac ttcaggaaca 12660
agcgggcgct gctcgacgca ctggccgaag ccatgctggc ggagaatcat acgcattcgg 12720
tgccgagagc cgacgacgac tggcgctcat ttctgatcgg gaatgcccgc agcttcaggc 12780
aggcgctgct cgctaccgc gatggcgcg gcacccatgc cggcacgca cggggcgac 12840

cgcagatgga aacggccgac ggcagcttc gcttcctctg cgaggcgggt ttttcggccg 12900
gggacgccgt caatgcgctg atgacaatca gctacttcac tgttggggcc gtgcttgagg 12960
agcaggccgg cgacagcga ggcggcgagc gggcgggcac cgttgaacag gctccgctct 13020
cgccgctgtt gcgggccgag atagacgcct tcgacgaagc cggtcgggac gcagcgcttcg 13080
agcagggact cgcggtgatt gtcgatggat tggcgaaaag gaggctcgtt gtcaggaacg 13140
ttgaaggacc gagaaagggt gacgattgat caggaccgct gccggagcgc aaccactca 13200
ctacagcaga gccatgtaga caacatcccc tcccccttc caccgcgtca gacgcccgta 13260
gcagcccgtt acgggctttt tcatgccctg ccctagcgtc caagcctcac ggccgcgctc 13320
ggcctctctg ggggccttct ggcgctcttc cgcttcctcg ctactgact cgctgcgctc 13380
ggtcgcttcg ctgcggcgag cggatcagc tcaactcaaag gcggtataac ggttatccac 13440
agaatcagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 13500
ccgtaaaaag gccgcgttgc tggcggtttt ccataggctc cgccccctg acgagcatca 13560
caaaaatcga cgctcaagtc agagggtggcg aaaccgcaca ggactataaa gataccaggc 13620
gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 13680
cctgtccgcc tttctccctt cgggaagcgt ggcgcttttc cgctgcataa ccctgcttcg 13740
gggtcattat agcgattttt tcggtatata catccttttt cgacagatat acaggatttt 13800
gccaaagggt tcgtgtagac tttccttggg gtatccaacg gcgtcagccg ggcaggatag 13860
gtgaagtagg cccacccgcg agcgggtgtt ccttcttcac tgtcccttat tcgcacctgg 13920
cgggtgctcaa cgggaatcct gctctgcgag gctggccggc taccgccggc gtaacagatg 13980
agggcaagcg gatggctgat gaaaccaagc caaccaggaa gggcagcca cctatcaagg 14040
tgtactgcct tccagacgaa cgaagagcga ttgaggaaaa ggcggcgggc gccggcatga 14100

gcctgtcggc ctacctgctg gccgtcggcc agggctacaa aatcacgggc gtcgtggact 14160
atgagcacgt ccgcgagctg gcccgcattca atggcgacct gggccgcctg ggcggcctgc 14220
tgaaactctg gctcaccgac gaccgcgcga cggcgcggtt cggatgatgcc acgatcctcg 14280
ccctgctggc gaagatcgaa gagaagcagg acgagcttgg caaggatcatg atgggcgtgg 14340
tccgcccag ggcagagcca tgactttttt agccgctaaa acggccgggg ggtgcgcgtg 14400
attgccaagc acgtcccat gcgctccatc aagaagagcg acttcgcgga gctggtgaag 14460
tacatcaccg acgagcaagg caagaccgag cgcctttgcg acgctcaccg ggctggttgc 14520
cctcgccgct gggctggcgg ccgtctatgg cctgcaaac gcgccagaaa cgccgtcgaa 14580
gccgtgtgcg agacaccgcg gccgcggcg ttgtggatac ctgcgcgaaa acttggccct 14640
cactgacaga tgaggggcgg acgttgacac ttgaggggccc gactcaccg gcgcggcggtt 14700
gacagatgag gggcaggctc gatttcggcc ggcgacgtgg agctggccag cctcgcaaatt 14760
cggcgaaaac gcctgatttt acgcgagttt cccacagatg atgtggacaa gcctggggat 14820
aagtgcctg cggtattgac acttgagggg cgcgactact gacagatgag gggcgcgac 14880
cttgacactt gaggggcaga gtgctgacag atgaggggcg cacctattga catttgaggg 14940
gctgtccaca ggcagaaaat ccagcatttg caagggtttc cgcccgtttt tcggccaccg 15000
ctaacctgtc ttttaacctg cttttaaac aatatttata aaccttgttt ttaaccaggg 15060
ctgcgccctg tgccgctgac cgcgcacgcc gaaggggggt gccccccctt ctgaaccct 15120
cccgcccgcc taacgcgggc ctcccatccc ccagggggt gcgcccctcg gccgcgaacg 15180
gcctcaccac aaaaatggca gcgctggcag tccttgccat tgccgggatc ggggcagtaa 15240
cgggatgggc gatcagcccg agcgcgacgc ccggaagcat tgacgtgccg caggtgctgg 15300
catcgacatt cagcgaccag gtgccgggca gtgagggcgg cggcctgggt ggcggcctgc 15360

ccttcacttc ggccgtcggg gcattcacgg acttcatggc ggggccggca atttttacct 15420
tgggcattct tggcatagt gtcgcgggtg ccgtgctcgt gttcgggggt gcgataaacc 15480
cagcgaacca tttgaggtga taggtaagat tataccgagg tatgaaaacg agaattggac 15540
ctttacagaa ttactctatg aagcgccata tttaaaaagc taccaagacg aagaggatga 15600
agaggatgag gaggcagatt gcottgaata tattgacaat actgataaga taatatatct 15660
tttatataga agatatcgcc gtatgtaagg atttcagggg gcaaggcata ggcagcgcgc 15720
ttatcaatat atctatagaa tgggcaaagc ataaaaactt gcatggacta atgcttgaaa 15780
cccaggacaa taaccttata gcttgtaa at tctatcataa ttgggtaatg actccaactt 15840
attgatagt ttttatgttc agataatgcc cgatgacttt gtcatgcagc tccaccgatt 15900
ttgagaacga cagcgacttc cgtcccagcc gtgccagggt ctgcctcaga ttcaggttat 15960
gccgctcaat tcgctgcgta tategettgc tgattacgtg cagctttccc ttcaggcggg 16020
attcatacag cggccagcca tccgtcatcc atatcaccac gtcaaagggt gacagcaggc 16080
tcataagacg cccagcgtc gccatagtgc gttcaccgaa tacgtgcgca acaaccgtct 16140
tccggagact gtcatacgcg taaaacagcc agcgctggcg cgatttagcc ccgacatagc 16200
cccactgttc gtccatttcc ggcagacga tgacgtcact gcccggtgt atgcgcgagg 16260
ttaccgactg cggcctgagt tttttaagt acgtaaaatc gtgttgaggc caacgcccat 16320
aatgcgggct gttgccggc atccaacgcc attcatggcc atatcaatga ttttctggtg 16380
cgtaccgggt tgagaagcgg tgtaagtga ctgcagttgc catgttttac ggcagtgaga 16440
gcagagatag cgctgatgtc cggcggtgct tttgccgta cgcaccaccc cgtcagtagc 16500
tgaacaggag ggacagctga tagacacaga agccactgga gcacctcaaa aacaccatca 16560
tacactaaat cagtaagttg gcagcatcac ccataattgt gggtttcaaaa tcggctccgt 16620

cgatactatg ttatacgcca actttgaaaa caactttgaa aaagctgttt tctggatatt 16680
aaggtttttag aatgcaagga acagtgaatt ggagttcgtc ttgttataat tagcttcttg 16740
gggtatcttt aaatactgta gaaaagagga aggaaataat aaatggctaa aatgagaata 16800
tcaccggaat tgaaaaaact gatcgaaaaa taccgctgcg taaaagatac ggaaggaatg 16860
tctcctgcta aggtatataa gctgggtggga gaaaatgaaa acctatattt aaaaatgacg 16920
gacagccggt ataaaggac cacctatgat gtggaacggg aaaaggacat gatgctatgg 16980
ctggaaggaa agctgcctgt tccaaaggtc ctgcactttg aacggcatga tggctggagc 17040
aatctgctca tgagtgaggc cgatggcgtc ctttgctcgg aagagtatga agatgaacaa 17100
agccctgaaa agattatcga gctgtatgcg gagtgcata ggctctttca ctccatcgac 17160
atatcggatt gtccctatac gaatagctta gacagccgct tagccgaatt ggattactta 17220
ctgaataacg atctggccga tgtggattgc gaaaactggg aagaagacac tccatttaaa 17280
gatccgcgcg agctgtatga ttttttaaag acggaaaagc ccgaagagga acttgtcttt 17340
tcccacggcg acctgggaga cagcaacatc tttgtgaaag atggcaaagt aagtggcttt 17400
attgatcttg ggagaagcgg cagggcggac aagtggatatg acattgcctt ctgcgtccgg 17460
tcgatcaggg aggatatcgg ggaagaacag tatgtcgagc tattttttga cttactgggg 17520
atcaagcctg attgggagaa aataaaatat tatattttac tggatgaatt gtttttagtac 17580
ctagatgtgg cgcaacgatg ccggcgacaa gcaggagcgc accgacttct tccgcatcaa 17640
gtgttttggc tctcaggcgg agggccacgg caagtatttg ggcaaggggt cgctggatatt 17700
cgtgcagggc aagattcgga ataccaagta cgagaaggac ggccagacgg tctacgggac 17760
cgacttcatt gccgataagg tggattatct ggacaccaag gcaccaggcg ggtcaaataca 17820
ggaataaggg cacattgccc cggcgtgagt cggggcaatc ccgcaaggag ggtgaatgaa 17880

tccggacgttt gaccggaagg catacaggca agaactgac gacgcgggggt tttccgccga 17940
ggatgccgaa accatcgcaa gccgcaccgt catgcgtgcg ccccgcgaaa cttccagtc 18000
cgtcggctcg atggtccagc aagctacggc caagatcgag cgcgacagcg tgcaactggc 18060
tccccctgcc ctgcccgcgc catcggccgc cgtggagcgt tcgcgtcgtc tcgaacagga 18120
ggcggcaggt ttggcgaagt cgatgaccat cgacacgcga ggaactatga cgaccaagaa 18180
gcgaaaaacc gccggcgagg acctggcaaa acaggtcagc gaggccaagc aggccgcgtt 18240
gctgaaacac acgaagcagc agatcaagga aatgcagctt tccttgttcg atattgcgcc 18300
gtggccggac acgatgcgag cgatgccaaa cgacacggcc cgctctgccc tgttcaccac 18360
gcgcaacaag aaaatcccgc gcgaggcgt gcaaaacaag gtcattttcc acgtcaacaa 18420
ggacgtgaag atcacctaca ccggcgtcga gctgcggggc gacgatgacg aactgggtgtg 18480
gcagcaggtg ttggagtacg cgaagcgcac ccctatcggc gagccgatca cttcacgtt 18540
ctacgagctt tgccaggacc tgggctggtc gatcaatggc cgggtattaca cgaaggccga 18600
ggaatgcctg tcgcgcctac aggcgacggc gatgggcttc acgtccgacc gcgttgggca 18660
cctggaatcg gtgtcgtgc tgcaccgctt ccgcgtcctg gaccgtggca agaaaacgtc 18720
ccgttgccag gtctgatcg acgaggaaat cgctcgtgctg tttgctggcg accactacac 18780
gaaattcata tgggagaagt accgcaagct gtcgccgacg gcccgacgga tgttcgacta 18840
tttcagctcg caccgggagc cgtaccgcgt caagctggaa accttccgcc tcatgtgcgg 18900
atcggattcc acccgctga agaagtggcg cgagcaggtc ggcaagcct gcgaagagtt 18960
gcgaggcagc ggcttgggtg aacacgcctg ggtcaatgat gacctggtgc attgcaaacg 19020
ctagggcctt gtggggtcag ttccggctgg ggggttcagca gccagcgctt tactggcatt 19080
tcaggaacaa gcggggcactg ctcgacgcac ttgcttcgct cagtatcgct cgggacgcac 19140

ggcgcgctct acgaactgcc gataaacaga ggattaaaat tgacaattgt gattaaggct 19200
cagattcgac ggcttgagc ggccgacgtg caggatttcc gcgagatccg attgtcggcc 19260
ctgaagaaag ctccagagat gttcgggtcc gtttacgagc acgaggagaa aaagcccatg 19320
gaggcgttcg ctgaacggtt gcgagatgcc gtggcattcg ggcctacat cgacggcgag 19380
atcattgggc tgtcgggtctt caaacaggag gacggcccca aggacgtca caaggcgcat 19440
ctgtccggcg ttttcgtgga gcccgaaacag cgaggccgag gggtcgccgg tatgctgctg 19500
cgggcgttgc cggggggttt attgctcgtg atgatcgtcc gacagattcc aacgggaatc 19560
tggtggatgc gcatcttcat cctcggcgca cttaatatct cgctattctg gagcttggtg 19620
tttatttcgg tctaccgcct gccgggcggg gtcggcgca cggtaggcgc tgtgcagccg 19680
ctgatggtcg tgttcatctc tgccgctctg ctaggtagcc cgatacgatt gatggcggtc 19740
ctgggggcta tttgcggaac tgcgggcgtg gcgctgttg tgttgacacc aaacgcagcg 19800
ctagatcctg tcggcgtcg agcgggcctg gcggggcgcg tttccatggc gttcggaacc 19860
gtgctgacct gcaagtggca acctcccggt cctctgctca cctttaccgc ctggcaactg 19920
gcggccggag gacttctgct cgttccagta gctttagtgt ttgatccgcc aatcccgatg 19980
cctacaggaa ccaatgttct cggcctggcg tggctcggcc tgatcggagc gggtttaacc 20040
tacttccttt ggttccgggg gatctcgcga ctcgaaacta cagttgtttc cttactgggc 20100
tttctcagcc ccagatctgg ggtcgatcag ccgggggatgc atcaggccga cagtcggaac 20160
ttcgggtccc cgacctgtac cattcgggtga gcaatggata ggggagttga tatcgtcaac 20220
gttcacttct aaagaaatag cgccactcag ctctctcagc ggctttatcc agcgatttcc 20280
tattatgtcg gcatagttct caagatcgac agcctgtcac ggttaagcga gaaatgaata 20340
agaaggctga taattcggat ctctgcgagg gagatgatat ttgatcacag gcagcaacgc 20400

tctgtcatcg ttacaatcaa catgctaccc tccgcgagat catccgtgtt tcaaaccggg 20460
cagcttagtt gccgttcttc cgaatagcat cggtaacatg agcaaagtct gccgccttac 20520
aacggctctc ccgctgacgc cgtcccggac tgatgggctg cctgtatcga gtggtgattt 20580
tgtgccgagc tgcgggtcgg ggagctgttg gctggctggg ggcaggatat attgtgggtg 20640
aaacaaattg acgcttagac aacttaataa cacattgcgg acgtttttaa tgtactgggg 20700
tggtttttct tttcaccagt gagacgggca acagctgatt gcccttcacc gcctggccct 20760
gagagagttg cagcaagcgg tccacgctgg tttgccccag caggcgaaaa tcctgtttga 20820
tggtgggtcc gaaatcggca aaatccctta taaatcaaaa gaatagcccg agatagggtt 20880
gagtgttggt ccagtttgga acaagagtcc actattaaag aacgtggact ccaacgtcaa 20940
agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac ccaaataag 21000
ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga gccccgatt 21060
tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga aagcgaaagg 21120
agcggggcgcc attcaggctg cgcaactgtt ggggaaggcg atcgggtgcgg gcctcttcgc 21180
tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag 21240
ggttttccca gtcacgacgt tgtaaaacga cggccagtga attcgagctc ggtaccggg 21300

<210> 47

<211> 17756

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>

<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>

<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 47

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaaggg ttcgcccgt	480
tttctggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaac cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc	660
tgggcgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgcca cgcccgaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg	840

ggtggcgggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcgggggccc 900
gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140
agataatata tcttttatat agaagatata gccgtatgta aggatttcag ggggcaaggc 1200
ataggcagcg cgcttatcaa tatactata gaatgggcaa agcataaaaa cttgcatgga 1260
ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcccctg agttttttta gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta tttaagggtt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatggtcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgagg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggc acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatcgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcg cttacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggctg tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgtatt 5100
ctggagcttg ttgtttatth cggtctaccg cctgcggggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtccctggggg ctattttcgg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcctggcaa ctggcgggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggctgat cagccgggga tgcacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt totcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcgc agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacgggt ctcccgtga cgcgcgtccg gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggccacgc tggtttgccc cagcaggcga 6300
aaatcctggt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900
gtcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320
tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaa gggcggttgca caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaate 7920
tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga cccagagggg catgacttga gcctaaaate cgccgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtacccggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaaag gccgctcagg agctggccag cggcgagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640

tcagtccttg gtaggcagct ttgccccgctc tgtccgcccgtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttgttgc catatcttcc tgcctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctcggtgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatgggtt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtc gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg tcgagcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtgggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc a gactcgtcc 9960
gaggggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcggaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacggcag tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagegc ccaactccaca tctccactcg acctgcaggc atgcaagctt cattttgctt 10800
tgtaaatttc tggtaactgc caccaagaaa tatgaggata ttcgtgatgt tcctcgtggt 10860
agccaaaatg atagcacgtg ataaatgacc accaaatagg acggctaatt gtttgggcac 10920

aatgaggctg aacataaccc cctattgggt cactatgggg taaaaaagta caaaataga 10980
ataattgtaa tgaacttaaa agcgagggtg gcacccaaaa gtaagttaga ttatcacttg 11040
ggatatggag tatgtattta gcaaagttat aaataatagt caacgcaatt atttgcccc 11100
aactccagta acctttcata aaatgaaaat accaagcaaa gaaactttgg tgtttaccat 11160
tgtgaaaatc cgggtctatt gagcttgctg gattgtgggt gtgtaaccaa tgttttttca 11220
atagtttttg atatggtaaa agaccataaa gggatagggt caatgttcca atcaaatgat 11280
taatcttggg gttttgggga aatactacgc catgcatggc atcatgagat gtaataaata 11340
atcccgtata taaaatggt tgccatagta taacaggcaa taacatccaa aatttttagct 11400
ttgagatgtc aagggaagt aataaactca ggctaattgac ccatgcgcta acaatgacaa 11460
tagcaatgaa aagcccctta aactgagatt tacttctcag tactggagtc agttttgctt 11520
gatgactgag tggttgttct aactggatca tttctaaaga gaaggtggaa caatgttagc 11580
ataattgtgc ttgagtgagg actttgaggg taggtacata cttgataaag ttaatgatta 11640
aagagaaaaa aaaagttttg gttcaaagca gaaattgttt tttaaactga ttggtgagaa 11700
aatttttttc tgtttcgca tcaccaaagc cacctcagga atggtcacaa attattggtc 11760
tgattggacc ataagcatac aaaaagttca ttgaagtata cttagtgggt tattagactt 11820
ttatcgtttt ctaacgcgaa tcagcaatgt ttcttgtttg atttactgct tgcttttagat 11880
catttttgtc tgaaatatta tgcatttggt caaagcggcc ttgttttctt ttctttcatg 11940
cttaaacacg ttgtttattc catatattac tttgaatatg catcacgca aagcggaagt 12000
gcaaaataac aaagaacctc tttgggttac acgatcaact gctattgtga aaaaaatttc 12060
tttttgaaaa tttttggaat aatatctctt gcaaaaaaga aattttgtat atttagtagc 12120
atcaagaaca aatgaaagaa gtgtgggata acaagaatac atcatcttta gacaaaagta 12180

cgagaaaaat ctaataagtt gttatagagg tctttgtttt ctttgtgttt atagacagtt 12240
atttagagtt tgaaaagtgt ctctaagtgt tcttttttta ttattattat ttcaaagtgt 12300
atgtaatata gctaaagcta tagatttgac attttttcta aatataaaat ttcagtcaac 12360
agaaataaat gacacgagtt ctttttctct ctctcaatcc tgttgatcat caatctttga 12420
tgtcgtttta aaacaaatga atggcattta gttccttagg tgtcactcac atcttgttga 12480
ccagaaaatc cttattcgcc ctcaaactctg ctttattcct ttcatttgat ttgatgttta 12540
agtaatgcaa gcaaacaaaa aagaaacctt tcttgcaaag acaaaagaat tgttttcaga 12600
ggaaagcaac tcgttgatcat tttttaagga tttagactta taatcgacac catagtttgt 12660
ccgttacatt ttttattgtc gttttctgat ttccttttaa tctttaagca aaatcaatat 12720
taacttatct tgtcttccaa taaaaaatgg ataccaataa caataaatcc ttcacaaaga 12780
aaaaaaaaaa aaactcgaaa aaagcttggc gtaatcatgg tcatagctgt ttcctgtgtg 12840
aaattgttat ccgctcaciaa ttccacacia catacgagcc ggaagcataa agtgtaaagc 12900
ctgggggtgcc taatgagtga gctaactcac attaatgctg ttgctgtcac tgcccgtttt 12960
ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgct cggggagagg 13020
cggtttgct attgggcaa agacaaaagg gcgacattca accgattgag ggaggggaagg 13080
taaattattga cggaaattat tcattaaagg tgaattatca ccgtcaccga cttgagccat 13140
ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa ggccggaaac 13200
gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat caagtttgcc 13260
tttagcgtca gactgtagct cgttttctatc ggcattttctg gtcataagccc cttattagc 13320
gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg aaccgcctcc 13380
ctcagagccg ccaccctcag aaccgccacc ctcagagcca ccaccctcag agccgccacc 13440

agaaccacca ccagagccgc cgccagcatt gacaggaggc ccgatctagt aacatagatg 13500

acaccgcgcg cgataattta tcctagtttg cgcgctatat ttgttttct atcgcgtatt 13560

aaatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac gtcatgcatt 13620

acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata atcatcgcaa 13680

gaccggcaac aggattcaat cttaagaaac ttatttgcca aatgtttgaa cgatcgggga 13740

tcacccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag atatcgcggt 13800

gcaccccggt cttgcctggg cagtcgccgc cgacgccgtt gatgtggacg ccgggcccga 13860

tcataattgtc gtcaggatc gtggcggtgt gcttgtcggc cgttgctgtc gtaatgatat 13920

cggcaccttc gaccgcctgt tccgcagaga tcccgtaggc gaagaactcc agcatgagat 13980

ccccgcgctg gaggatcatc cagccggcgt ccgggaaaac gattccgaag cccaaccttt 14040

catagaaggc ggcgggtggaa tcgaaatctc gtgatggcag gttgggcgtc gcttggtcgg 14100

tcatttcgaa cccagagtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat 14160

gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc 14220

gccaaagctct tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac 14280

accagccgg ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg 14340

caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc gcgccttgag 14400

cctggcgaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc 14460

gacaagaccg gcttccatcc gactacgtgc tcgctcgatg cgatgtttcg cttggtggtc 14520

gaatgggcag gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga 14580

tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa 14640

tagcagccag tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc 14700

cgtcgtggcc agccacgata gccgcgctgc ctgcgtctgc agttcattca gggcaccgga 14760
caggteggtc ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc 14820
atcagagcag ccgattgtct gttgtgcccc gtcatagccg aatagcctct ccaccaagc 14880
ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgatc cagatccggc 14940
gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag gggaatttat 15000
ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg accttaggcg 15060
acttttgaac gcgcaataat ggtttctgac gtatgtgctt agtcattaa actccagaaa 15120
cccgcggtg agtggctcct tcaacgttgc gggtctgtca gttccaaacg taaaacggct 15180
tgtcccgct catcggcggg ggtcataacg tgactccctt aattctccgc tcatgatcag 15240
attgtcgttt cccgccttca gtttaaacta tcagtgttg acaggatata ttggcgggta 15300
aacctaagag aaaagagcgt ttattagaat aatcgatat ttaaaagggc gtgaaaaggc 15360
ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga tctggcgccg 15420
gccagcgaga cgagcaagat tggccgcccgc ccgaaacgat ccgacagcgc gccagcaca 15480
ggtgcgagcagg caaattgcac caacgcatac agcgccagca gaatgccata gtgggcggtg 15540
acgtcgttcg agtgaaccag atcgcgagcagg agggccggca gcaccggcat aatcaggccg 15600
atgccgacag cgtcgagcgc gacagtgtc agaattacga tcaggggtat gttgggtttc 15660
acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggtt ctttatcact 15720
gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg acaaagttgc 15780
agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggtcggc gtagacggtc 15840
tgacgacacg caaactggcg gaacggttgg gggttcagca gccggcgctt tactggcact 15900
tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc catgctggcg gagaatcata 15960

cgcatcgggt gccgagagcc gacgacgact ggcgctcatt tctgatcggg aatgcccgca 16020
gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg catccatgcc ggcacgcgac 16080
cgggcgccacc gcagatggaa acggccgacg cgcagcttcg cttcctctgc gagggcgggtt 16140
tttcggcccg ggacgccgtc aatgcgctga tgacaatcag ctacttcact gttggggccg 16200
tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc gttgaacagg 16260
ctccgctctc gccgctgttg cgggccgcca tagacgcctt cgacgaagcc ggtccggacg 16320
cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg aggctcgttg 16380
tcaggaacgt tgaaggaccg agaaagggcg acgattgatc aggaccgctg ccggagcgca 16440
accactcac tacagcagag ccatgtagac aacatccctt ccccttttcc accgcgtcag 16500
acgcccgtag cagcccgcta cgggcttttt catgccctgc cctagcgtcc aagcctcacg 16560
gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttcctcgc tactgactc 16620
gctgcgctcg gtcgttcggc tgccggcgagc ggtatcagct cactcaaagg cggtaatacg 16680
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 16740
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 16800
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgacag gactataaag 16860
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 16920
taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgcttttcc gctgcataac 16980
cctgcttcgg ggtcattata gcgatttttt cggtatatcc atcctttttc gcacgatata 17040
caggattttg ccaaaggggt cgtgtagact ttccttggtg tatccaacgg cgtcagccgg 17100
gcaggatagg tgaagtaggc ccacccgcca gcgggtgttc cttcttctact gtcccttatt 17160
cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg ctggccgggt accgccggcg 17220

taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag ggcagcccac 17280
ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag gcggcggcgg 17340
ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa atcacgggcg 17400
tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa tggcgacctg ggccgcctgg 17460
gcggcctgct gaaactctgg ctcaccgacg acccgcgcac ggcgcggttc ggtgatgcca 17520
cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc aaggtcatga 17580
tgggcgtggt ccgcccgagg gcagagccat gactttttta gccgctaaaa cggccggggg 17640
gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga cttcgcgag 17700
ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcctttgcga cgctca 17756

<210> 48

<211> 17118

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 48

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc ggggccgccg gcgttggtgga tacctcgcg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaaggg ttcgcccgt	480
ttttcgcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccc	600
cttctcgaa cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgccc	660
tcggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgga cgcccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg	840
ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140
agataatata tcttttatat agaagatc gccgtatgta aggatttcag ggggcaaggc	1200

ataggcagcg cgcttatcaa tatactata gaatgggcaa agcataaaaa cttgcatgga 1260
ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100
aatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttggggatat tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggc attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cgggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360
gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggt gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcatgtcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg ttatttgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata ttctgctatt 5100
ctggagcttg ttgtttatatt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtcctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgccctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc ttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcattcaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcgg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cggcgtccc gactgatggg ctgcctgtat 6060
cgagtgggta ttttggtccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggccacgc tggtttgccc cagcaggcga 6300
aaatcctggt tgatgggtgg tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
caccctaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900
gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320
tcctttggct ttcaatatcc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaa gggcggttgca caatgcaagt caaacggtgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920
tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga cccagagggc catgacttga gcctaaaadc cgccgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgtatga aaccggaag gccgctcagg agctggccag cggcgagac 8580
cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggctggtgc gtcagtcaca catttgttgc catattttcc tgcctctccc accagctgct 8760

ctttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctctgtcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgccgct gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaacgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcacaaaca gcttgacgaa tctggatata agatcggttg tgatgatgtc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacggtc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctggggt cgcaaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctctgttact 10740
gtgtaagcgc ccaactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800
tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160
cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac tctaaatctc gatgaccga caaaaagag cgaacaaaga aataatcttg 11280

tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400

tatgatccag ttagaacaac cactcagtca tcaagcaaaa ctgactccag tactgagaag 11460

taaatctcag tttaaggggc ttttcattgc tattgtcatt gttagcgcac gggtcattag 11520

cctgagttta ttactttccc ttgacatctc aaagctaaaa ttttggtatg tattgcctgt 11580

tatactatgg caaacatttt tatatacggg attatttatt acatctcatg atgccatgca 11640

tggcgtagta tttcccaaaa acaccaagat taatcatttg attggaacat tgaccctatc 11700

cctttatggt cttttaccat atcaaaaact attgaaaaaa cattgggttac accaccacaa 11760

tccagcaagc tcaatagacc cggattttca caatggtaaa caccaaagtt tctttgcttg 11820

gtattttcat tttatgaaag gttactggag ttgggggcaa ataattgcgt tgactattat 11880

ttataacttt gctaaataca tactccatat cccaagtgat aatctaactt acttttgggt 11940

gctaccctcg cttttaagtt cattacaatt attctatttt ggtacttttt taccctatag 12000

tgaaccaata ggggggttatg ttcagcctca ttgtgcccac acaattagcc gtcctatttg 12060

gtggtcattt atcacgtgct atcatttttg ctaccacgag gaacatcacg aatctcctca 12120

tatttcttgg tggcagttac cagaaattta caaagcaaaa tagaagcttg gcgtaatcat 12180

ggcatagct gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacgag 12240

ccggaagcat aaagtgtaaa gcctgggggtg cctaattgag gagctaactc acattaattg 12300

cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 12360

tcggccaaag cgcgggggaga ggcgggtttgc gtattgggcc aaagacaaaa gggcgacatt 12420

caaccgattg agggaggggaa ggtaaatatt gacggaaatt attcattaaa ggtgaattat 12480

caccgtcacc gacttgagcc atttggaat tagagccagc aaaatcacca gtagcaccat 12540

taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12600
tagcgacaga atcaagtttg cctttagcgt cagactgtag cgcgttttca tcggcatttt 12660
cggtcatagc ccccttatta gcgtttgcca tcttttcata atcaaaatca ccggaaccag 12720
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgcca ccctcagagc 12780
caccaccctc agagccgcca ccagaaccac caccagagcc gccgccagca ttgacaggag 12840
gcccgatcta gtaacataga tgacaccgcg cgcgataatt taccctagtt tgccgcgtat 12900
attttgtttt ctatcgcgta tttaatgtat aattgcggga ctctaatacat aaaaacccat 12960
ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13020
aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13080
caaagtgttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13140
gaacgcagca agatatcgcg gtgcatctcg gtcttgccctg ggcagtcgcc gccgacgccg 13200
ttgatgtgga cgccggggcc gatcatattg tcgctcagga tcgtggcggt gtgcttgctg 13260
gccgttgctg tcgtaatgat atcggcacct tcgaccgcct gttccgcaga gatcccgtag 13320
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccgga 13380
acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13440
aggttgggcg tcgcttggtc ggtcatttcg aacccagag tcccgctcag aagaactcgt 13500
caagaaggcg atagaaggcg atgcgtgcg aatcgggagc ggcgataccg taaagcacga 13560
ggaagcggtc agccattcg ccgccaagct cttcagcaat atcacgggta gccaacgcta 13620
tgtcctgata gcggtcgcc acaccagcc ggccacagtc gatgaatcca gaaaagcggc 13680
cattttccac catgatattc ggcaagcagg catcgccatg ggtcacgacg agatcatcgc 13740
cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13800

cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtacgt gctcgtcga 13860
tgcgatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 13920
gcattgcac agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 13980
cctgccccgg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga 14040
gcacagctgc gcaaggaacg cccgtcgtgg ccagccacga tagccgcgct gcctcgtcct 14100
gcagttcatt cagggcacccg gacaggtcgg tcttgacaaa aagaaccggg cgcccctgcg 14160
ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14220
cgaatagcct ctccacccaa ggggccggag aacctgcgtg caatccatct tgttcaatca 14280
tgcgaaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14340
ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatttgc 14400
tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14460
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14520
cagttccaaa cgtaaaacgg cttgtcccgc gtcacggcg ggggtcataa cgtgactccc 14580
ttaattctcc gtcatgatc agattgtcgt tccccgctt cagtttaaac tatcagtgtt 14640
tgacaggata tattggcggg taaacctaa agaaaagagc gtttattaga ataatcggat 14700
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaaccaca 14760
gggttcccca gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaacg 14820
atccgacagc gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag 14880
cagaatgcca tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg 14940
cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15000
gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15060

aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15120
gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15180
aacgaggtcg gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt gggggttcag 15240
cagccggcgc ttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15300
gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15360
tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15420
cgcattccatg ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgacagctt 15480
cgcttcctct gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc 15540
agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15600
cgcggcggca ccgttgaaca ggctccgctc tcgccgctgt tcggggccgc gatagacgcc 15660
ttcgacgaag ccggtccgga cgcagcgctc gagcaggggac tcgcggtgat tgtcgatgga 15720
ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15780
tcaggaccgc tgccgggagcg caaccctc actacagcag agccatgtag acaacatccc 15840
ctcccccttt ccaccgcgct agacgcccggt agcagccgc tacgggcttt ttcatgccct 15900
gccctagcgt ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt 15960
ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag 16020
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16080
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgctt ctggcgcttt 16140
tccataggct ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagagggtggc 16200
gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc ctcgctgcgct 16260
ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 16320

tggcgctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat 16380
ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16440
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt 16500
tccttcttca ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga 16560
ggctggccgg ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16620
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16680
attgaggaaa agggggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16740
cagggttaca aaatcacggg cgctgtggac tatgagcacg tccgcgagct ggcccgcatc 16800
aatggcgacc tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc 16860
acggcgcggt tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag 16920
gacgagcttg gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt 16980
tagccgctaa aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat 17040
caagaagagc gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17100
gcgcctttgc gacgctca 17118

<210> 49

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 49

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt	60
cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct	240
tcgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca	300
gccagggcca gccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggtttg ctttccaggc tgagactcta gcttggagca tagaggggtcc tttggctttc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggg cgttgcacia tgcaagtcaa acgttgctag cagttccagg	780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacaaaa gtcacacaac acaagctgta aggatttcgg	900

caaggctacg gaagacggag aagccacctt cagtggactc gaggaccatt taattctatt 960
tgtgtttgat cgagacctaa tacagccctt acaacgacca tcaaagtcgt atagctacca 1020
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080
ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140
gttgaccggg gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200
ggagtgacct agaggggtcat gacttgagcc taaaatccgc cgctccacc atttgtagaa 1260
aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320
gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380
ctctgaggtg cagtggatga ttattaatcc gggaccggcc gcccctccgc cccgaagtgg 1440
aaaggctggg gtgcccctcg ttgaccaaga atctattgca tcctcgga atagggagct 1500
tcctcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560
cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttcgatc tggtaaaaga 1620
ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680
aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctctg ctttgcccgg 1740
tgtatgaaac cggaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800
tggcagtcga cccatccggg gctctgcact cgacctgtg aggtccctca gtccctggta 1860
ggcagctttg ccccgctctgt ccgcccgggtg tgtcggcggg gttgacaagg tcgttgcgtc 1920
agtcacaat ttgttgccat attttctgc tctccccacc agctgctctt ttcttttctc 1980
tttcttttcc catcttcagt atattcatct tccatccaa gaacctttat tccccctaag 2040
taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100
gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctccgagggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcgctg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgtc tgacattggg 2400

gaattcagcg agagcctgac ctattgcac tcccgcctg cacaggggtg cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagcccg tcgcggaggc catggatgcg 2520

atcgtgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580

ggccaataca ctacatggcg tgatttcata tgccgcgattg ctgatcccca tgtgtatcac 2640

tggcaaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga ttccggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgcagg atcgccgcgg 2940

ctccgggctg atatgtccg cattgggtctt gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtccg atccggagcc 3060

gggactgtcg ggcgtacaca aatcgccgcg agaagcgcg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgccccagca ctcgctccgag ggcaaaggaa 3180

tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct ctccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540

tcaacggctcg tcaagagacc tacgagactg aggaatccgc tcttggctcc acgcgactat 3600

atatttgtct ctaattgtac tttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780

gatggtatac aatagtaacc atgcatgggt gcctagttaa tgctccgtaa cacccaatac 3840

gccggccgaa actttttttac aactctccta tgagtcgttt acccagaatg cacaggatca 3900

cttggttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020

agaaagttaa aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080

ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140

ggagtttgaa attttgtcct ccaggagcaa aaaaaagttt ttttttatac atgtttgtac 4200

acaagaatag ttaccaattt gctttggtct tacgtgctgc aagtttatat cgttttcaat 4260

ttctttgtct ttacattttc tttgtccttt atctttcctc atttagtctt tgggagaatt 4320

aggaaaaggg agcggaaagg taagaaatgc ttgcgtattt tactaattcg gcaaacatcc 4380

aatttggcaa acagcagcct gtgcaacgct ctcgagatga cagtatcttt gattacactc 4440

taaatctcga tgacccgacc aaaaagagcg aacaaagaaa taatcttgtg cattogaata 4500

tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560

aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtcgaa 4620

gctgcagtca atcagcgtca aggcccgcgc cggtgaacta gcccgcgaca tcacgcggcc 4680

caaagtctgc ctgcatgctc agcggtgctc gttagttcgg ctgcgagtgg cagcaccaca 4740
gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800
tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccggc gcaaacggga 4860
gcagctgtca taccaggctg ccgccattgc agcatcaatt ggctgtcag gcattgccat 4920
cttcgccacc tacctgagat ttgccatgca catgaccgtg ggcggcgag tgccatgggg 4980
tgaagtggct ggcactctcc tcttggtggt tgggtggcgcg ctccggcatgg agatgtatgc 5040
ccgctatgca cacaagcca tctggcatga gtcgcctctg ggctggctgc tgcacaagag 5100
ccaccacaca cctcgactg gacccttga agccaacgac ttgtttgcaa tcatcaatgg 5160
actgcccgcc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tgggggcggc 5220
ctgctttgga gcggggctgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280
tggcctgggtg cacaggcgct ttcccaccgg gcccatcgct ggcctgccct acatgaagcg 5340
cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctgggggtat 5400
gttcttgggt ccacaggagc tgcagcacat tccagggtcg gcggaggagg tggagcgact 5460
ggctctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520
atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaatac agcaaggcaa 5580
cgcccgtagc ctaggtgatc aaagactgtt gggcttgtct ctgaagcttg taggaaaggc 5640
agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700
cgggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760
gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaaga ttgctctgtt 5820
gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880
gattaagagt ttttttggtg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

aaagagacaa cacatacaat ataaaatcat tgtcttttaa tgtaattta gtagagtgt 6000

aagcctgcat tttttttgta cgcataaaca atgaattcac cccgcttctg gtttttaa 6060

aattatgtca aactagggaa aattcttttt tttctcttcg ttcttttttt ggcttggtgt 6120

ggagtcacag gcttgtcttc agattgatag aggttgata cactcaacag agcaatcttg 6180

gcacgttcgc ttccttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt 6240

acagatccaa ttgcagcata tacttgga gacgctcgt caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcctacaagc 6360

ttcagagaca agcccaacag tctttgatca cctaggctac gggcggttgcc ttgctgtatt 6420

tggtctgttt ctctttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480

gagccatact acagcataaa aggatacgtt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac aaaaaaaaaag aaatttaaga tgagtaggac ttccattctc tcaaaaattt 6600

tattcaatcc ataaatgaat tatttttgga caaaaaagaa agattatgcc tgattttctc 6660

tatttttttt ttttttacia ctccaccaat actttctagc ccagcttggc gtaatcatgg 6720

tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc 6780

ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg 6840

ttgcgctcac tgcccgtttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900

ggccaacgcg cggggagagg cggtttgcgt attgggcaa agacaaaagg gcgacattca 6960

accgattgag ggagggaagg taaatattga cggaaattat tcattaaagg tgaattatca 7020

ccgtcacga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140

gcgacagaat caagtttgcc ttagcgctca gactgtagcg cgttttcatc ggcattttcg 7200

gtcatagccc ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag 7260
ccaccaccgg aaccgcctcc ctgagagccg ccaccctcag aaccgccacc ctgagagcca 7320
ccaccctcag agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggg 7380
ccgatctagt aacatagatg acaccgcgcg cgataattta tcttagtttg cgcgctatat 7440
tttgttttct atcgcgtatt aaatgtataa ttgcgggact ctaatcataa aaacccatct 7500
cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560
ttatatgata atcatcgcaa gaccggcaac aggattcaat cttaagaaac ttatttgcca 7620
aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680
acgcagcaag atatcgcggg gcactctcgg cttgcctggg cagtcgccgc cgacgccgtt 7740
gatgtggacg ccggggccga tcatattgtc gctcaggatc gtggcgttgt gcttgtcggc 7800
cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc 7860
gaagaactcc agcatgagat cccgcgcgtg gaggatcatc cagccggcgt cccggaaaac 7920
gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980
gttgggcgtc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040
agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100
aagcggtcag cccattcgcc gccaaagctct tcagcaatat cacgggtagc caacgctatg 8160
tctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220
ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg 8280
tcgggcatgc gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct 8340
tcgtccagat catcctgatc gacaagaccg gcttccatcc gactacgtgc tcgctcgatg 8400
cgatgtttcg cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460

attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520

tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc 8580

acagctgcgc aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc 8640

agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct 8700

gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg 8760

aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820

cgaaacgata cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880

ggataccgag ggggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940

gctgatagtg accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000

agtcattaa actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca 9060

gttccaaacg taaaacggct tgtcccgct catcgccggg ggtcataacg tgactccctt 9120

aattctccgc tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 9180

acaggatata ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat 9240

ttaaaagggc gtgaaaaggc ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300

gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgcccgc ccgaaacgat 9360

ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420

gaatgccata gtgggcgggtg acgtcgttcg agtgaaccag atcgcgcagg aggccccgca 9480

gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga 9540

tcaggggtat gttgggttct acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600

cgcgcggtat ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa 9720

cgagggtcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780
gccggcgctt tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc 9840
catgctggcg gagaatcata cgcattcggt gccgagagcc gacgacgact ggcgctcatt 9900
tctgatcggg aatgcccga gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg 9960
catccatgcc ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg 10020
cttcctctgc gaggcgggtt ttccggccgg ggacgccgtc aatgcgctga tgacaatcag 10080
ctacttcact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140
cggcggcacc gttgaacagg ctccgctctc gccgctgttg cgggcgcga tagacgcctt 10200
cgacgaagcc ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt 10260
ggcgaaaagg aggcctcgtt tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc 10320
aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380
ccccctttcc accgcgtcag acgcccgtag cagcccgcta cgggcttttt catgccctgc 10440
cctagegtcc aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500
gcttcctcgc tactgactc gctgcgctcg gtcgttcggc tgccggcagc ggtatcagct 10560
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740
aaccgcagag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 10800
cctgttccga cctgcgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg 10860
gcgcttttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggatatatc 10920
atcctttttc gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttggtg 10980

tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gggggtgttc 11040
cttcttcaact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100
ctggccgggt accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160
aaccaggaag ggcagccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11220
tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg cgtcggcca 11280
gggctacaaa atcacggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340
tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgac 11400
ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460
cgagcttggc aaggtcatga tgggcgtggt ccgcccagg gcagagccat gactttttta 11520
gccgctaaaa cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580
agaagagcga cttcggggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc 11640
gcctttgcga cgctcaccgg gctggttgcc ctcgccgtg ggtggcggc cgtctatggc 11700
cctgcaaacg cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt 11760
tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820
tgaggggccc actcaccgg cgcggcgttg acagatgagg ggcaggctcg atttcggccg 11880
gcgacgtgga gctggccagc ctcgcaaate ggcgaaaacg cctgatttta cgcgagtttc 11940
ccacagatga tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc 12000
gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060
tgaggggccc acctattgac atttgagggg ctgtccacag gcagaaaac cagcatttgc 12120
aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 12180
atatttataa accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg 12240

aaggggggtg ccccccttc tcgaaccctc ccggcccgt aacgcgggcc tcccatcccc 12300
ccagggggtg cggccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt 12360
ccttgccatt gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc 12420
cggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccggggcag 12480
tgagggcggc ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga 12540
cttcatggcg gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600
cgtgctcgtg ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt 12660
ataccgaggt atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat 12720
ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat 12780
attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840
tttcaggggg caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaaagca 12900
taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960
ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc 13020
gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg 13080
tgccaggtgc tgccctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140
gattacgtgc agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca 13200
tatcaccacg tcaaagggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg 13260
ttcaccgaat acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca 13320
gcgctggcgc gatttagccc cgacatagcc cactgttcg tccatttccg cgcagacgat 13380
gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga 13440
cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca 13500

ttcatggcca tatcaatgat tttctggtgc gtaccggggt gagaagcggg gtaagtgaac 13560
tgcagttgcc atgtttttacg gcagtgagag cagagatagc gctgatgtcc ggcggtgctt 13620
ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680
gccactggag cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc 13740
cataattgtg gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac 13800
aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860
gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920
ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980
accgctgcgt aaaagatacg gaaggaaatgt ctctgctaa ggtatataag ctggtgggag 14040
aaaatgaaaa cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg 14100
tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc 14160
tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc 14220
tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280
agtgcacag gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag 14340
acagccgctt agccgaattg gattacttac tgaataacga tctggccgat gtggattgag 14400
aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga 14460
cggaaaagcc cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520
ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580
agtggatatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt 14640
atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700
atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760

caggagcgca ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc 14820
aagtatttgg gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac 14880
gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940
gacaccaagg caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtgagtc 15000
ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060
gaactgatcg acgcgggggtt ttccgcccag gatgccgaaa ccatcgcaag ccgcaccgtc 15120
atgcgtgcgc cccgcgaaac cttccagtcc gtcggctcga tgggccagca agctacggcc 15180
aagatcgagc gcgacagcgt gcaactggct cccctgccc tgcccgcgc atcggccgcc 15240
gtggagcggt cgcgtcgtct cgaacaggag gcggcaggtt tggcgaagtc gatgaccatc 15300
gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa 15360
caggtcagcg aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420
atgcagcttt ccttggtcga tattgcgccg tggccggaca cgatcgagc gatgccaaac 15480
gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540
caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag 15600
ctgcggggccg acgatgacga actggtgtgg cagcaggtgt tggagtacgc gaagcgcacc 15660
cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggtcg 15720
atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780
atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgctcgtgct gcaccgcttc 15840
cgcgtcctgg accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc 15900
gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960
tcgccgacgg cccgacggat gttcgactat ttcagctcgc accgggagcc gtacccgctc 16020

aagctgga aa ccttccgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080
gagcaggctg gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg 16140
gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200
ggttcagcag ccagcgcctt actggcattt caggaacaag cgggcactgc tcgacgcact 16260
tgcttcgctc agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag 16320
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380
aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500
tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 16560
acggcccca aa ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc 16620
gaggccgagg ggtcgccggt atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga 16680
tgatcgtccg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac 16740
ttaatatctc gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccgggcgggg 16800
tcgcggcgac ggtaggcgct gtgcagccgc tgatggtcgt gttcatctct gccgctctgc 16860
taggtagccc gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg 16920
cgctgttggg gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 16980
cgggggcggg ttccatggcg ttcggaaccg tgctgacctg caagtggcaa cctcccgtgc 17040
ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 17100
ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggccctggcgt 17160
ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220
tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280

cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag 17340
caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400
ttcctcagcg gctttatcca gcgatttcct attatgtcgg catagtcttc aagatcgaca 17460
gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580
ccgcgagatc atccgtgttt caaaccoggc agcttagttg ccgttcttcc gaatagcatc 17640
ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact 17700
gatgggctgc ctgtatcgag tggtgathtt gtgccgagct gccggtcggg gagctgttgg 17760
ctggctgggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17820
acattgcgga cgtttttaat gtactgggggt ggtttttctt ttcaccagtg agacgggcaa 17880
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg 17940
ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18000
aaatcaaaag aatagcccgga gatagggttg agtggtgttc cagtttgga caagagtcca 18060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120
ccactacgtg aaccatcacc caaatcaagt tttttggggg cgaggtgccg taaagcacta 18180
aatcggaacc ctaaaggag ccccccgttt agagcttgac ggggaaagcc ggcgaacgtg 18240
gcgagaaagg aagggaagaa agcgaaagga gcggggcgcca ttcaggctgc gcaactgttg 18300
ggaaggcgga tcgggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcgagctcg gtaccggg 18449

<210> 50
<211> 18617
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 50
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgccg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgccgc gttgacagat gaggggacagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtgta cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatatatt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctogaac cctcccggcc cgctaacgcg ggccctcccat cccccaggg gctgcgcccc 660

tccggccgga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgcca cgcccgggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac cagggtgccg gcagtgaggg cggcggcctg 840

ggtggcggcc tgcccttcac ttcggccgctc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gaggggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaaggtata taagctgggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatatatt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatggtec agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcgggtgtcg tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggccccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcatgtcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttatct cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtctctgggg ctatttgcg aactgcgggc gtggcgctgt tgggtgttgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagegatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacgggtaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgccgtcccg gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgccg agctgccggg cggggagctg ttggctgggt ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300
aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcgggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcc 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaa ggcgcttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaata 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggg catgacttga gcctaaaata cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgctgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaatctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccgc 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtgcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcg gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300
ggccatggat gcgatcgctg cggccgatct tagccagacg agcggggttcg gccatttcgg 9360

accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggctgcc aacatcttct tctggaggcc 9660
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgtc agtccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtocaa gcagcaaaga 10140
gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatttg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800
tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160
cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac tctaaatctc gatgaccgga caaaaagag cgaacaaaga aataatcttg 11280
tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340
tggtatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400
tatgctgtcg aagctgcagt caatcagcgt caaggccgc cgcgttgaa tagcccgca 11460
catcacgagg cccaaagtct gcctgcatgc tcagcgggtc tcgttagttc ggctgcgagt 11520
ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgccggcg cgggcgatga 11580
gcacagcgcc gatgtagcac tccagcagct tgaccgggct atcgagagc gtcgtgccc 11640
gcgcaaacgg gagcagctgt cataccaggc tgccgccatt gcagcatcaa ttggcgtgtc 11700
aggcattgcc atcttcgcca cctacctgag atttgccatg cacatgaccg tgggcggcgc 11760
agtgccatgg ggtgaagtgg ctggcactct cctcttggtg gttggtggcg cgctcggcat 11820
ggagatgtat gcccgctatg cacacaaagc catctggcat gagtcgcctc tgggctggct 11880

gctgcacaag agccaccaca cacctcgcac tggacccttt gaagccaacg acttgtttgc 11940
aatcatcaat ggactgcccg ccatgctcct gtgtaccttt ggcttctggc tgcccaacgt 12000
cctggggggcg gcctgctttg gagcggggct gggcatcacg ctatacggca tggcatatat 12060
gtttgtacac gatggcctgg tgcacaggcg ctttcccacc gggcccatcg ctggcctgcc 12120
ctacatgaag cgcctgacag tggcccacca gctacaccac agcggcaagt acggtggcgc 12180
gccctgggggt atgttcttgg gtccacagga gctgcagcac attccaggtg cggcggagga 12240
gggtggagcga ctggtcctgg aactggactg gtccaagcgg tagaagcttg agattaaaat 12300
agataaggaa aagaaagtga aaagaaattc ggaagcatgg cacattcttc tttttataaa 12360
tacatgcctg actttctttt tccatcgata tgatatatgc atatgataga tatacaagca 12420
atcttcttca aggagtttga aattttgtcc tccaggagca aaaaaaagtt ttttttata 12480
catgtttgta cacaagaata gttaccaatt tgctttggtc ttacgtgctg caagtttata 12540
tcgttttcaa tttctttgtc ttacatttt ctttgtcctt tatctttcct catttagtct 12600
ttggggagaat taggaaaagg gagcggaaag gtaagaaatg cttgcgtatt ttactaatc 12660
ggcaaacatc caatttggca aacagcagcc tgtgcaacgc tctcgagatg acagtatctt 12720
tgattacact ctaaattctg atgacccgac caaaaagagc gaacaaagaa ataattctgt 12780
gcattcgaat atgatggaag attttttccc ccttattcta aatgttgaca tagcgtgtat 12840
gttatataaa caaaaagaaa ttgtacaaac tttcttttct tctcttttta ttttatctct 12900
atgatccagt tagaacaacc actcagtcac caagcaaac tgactccagt actgagaagt 12960
aaatctcagt ttaaggggct tttcattgct attgtcattg ttagcgcacg ggtcattagc 13020
ctgagtttat tactttccct tgacatctca aagctaaaat tttggatggt attgcctggt 13080
atactatggc aaacattttt atatacggga ttatttatta catctcatga tgccatgcat 13140

ggcgtagtat ttccccaaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 13200
ctttatggtc ttttaccata tcaaaaacta ttgaaaaaac attgggttaca ccaccacaat 13260
ccagcaagct caatagaccc ggattttcac aatggtaaac accaaagttt ctttgcttgg 13320
tattttcatt ttatgaaagg ttactggagt tgggggcaaa taattgcgtt gactattatt 13380
tataactttg ctaaatacat actccatata ccaagtata atctaactta cttttgggtg 13440
ctaccctcgc ttttaagttc attacaatta ttctattttg gtactttttt accccatagt 13500
gaaccaatag ggggttatgt tcagcctcat tgtgcccaaa caattagccg tcctatttgg 13560
tggtcattta tcacgtgcta tcattttggc taccacgagg aacatcacga atatcctcat 13620
atttcttggg ggcagttacc agaaatttac aaagcaaaat agaagcttgg cgtaatcatg 13680
gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc 13740
cggaagcata aagtgtaaag cctgggggtgc ctaatgagt agctaactca cattaattgc 13800
gttgogctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc attaataat 13860
cggccaaacgc gcgggggagag gcggtttgcg tattgggcca aagacaaaag ggcgacattc 13920
aaccgattga gggaggggaag gtaaattattg acggaaatta ttcattaaag gtgaattatc 13980
accgtcaccg acttgagcca tttgggaatt agagccagca aatcaccag tagcaccatt 14040
accattagca aggccggaaa cgtcaccaat gaaaccatcg atagcagcac cgtaatcagt 14100
agcgacagaa tcaagtttgc ctttagcgtc agactgtagc gcgttttcat cggcattttc 14160
ggtcatagcc cccttattag cgtttgccat cttttcataa tcaaaatcac cggaaccaga 14220
gccaccaccg gaaccgcctc cctcagagcc gccaccctca gaaccgccac cctcagagcc 14280
accaccctca gagccgccac cagaaccacc accagagccg ccgccagcat tgacaggagg 14340
cccgatctag taacatagat gacaccgcgc gcgataattt atcctagttt gcgcgctata 14400

ttttgttttc tatcgcgatat taaatgtata attgcggggac tctaatacata aaaacccatc 14460
tcataaataa cgatcatgcat tacatgttaa ttattacatg cttaacgtaa ttcaacagaa 14520
attatatgat aatcatcgca agaccggcaa caggattcaa tottaagaaa ctttattgcc 14580
aaatgtttga acgatcgggg atcatccggg tctgtggcgg gaactccacg aaaatatccg 14640
aacgcagcaa gatatcgcg tgcatctcgg tcttgccctgg gcagtcgccg ccgacgccgt 14700
tgatgtggac gccggggccc atcatattgt cgctcaggat cgtggcgctg tgcttgctcg 14760
ccgttgctgt cgtaatgata tcggcacctt cgaccgcctg ttccgcagag atcccgtggg 14820
cgaagaactc cagcatgaga tccccgcgct ggaggatcat ccagccggcg tcccggaaaa 14880
cgattccgaa gcccaacctt tcatagaagg cggcgggtgga atcgaaatct cgtgatggca 14940
ggttggggcg cgttggtcg gtcatttcga accccagagt cccgctcaga agaactcgtc 15000
aagaaggcga tagaaggcga tgcgctgcga atcgggagcg gcgataccgt aaagcacgag 15060
gaagcgggtca gcccatcgc cgccaagctc ttcagcaata tcacgggtag ccaacgctat 15120
gtcctgatag cgggtccgcca caccagccg gccacagtcg atgaatccag aaaagcggcc 15180
attttccacc atgatattcg gcaagcaggc atcgccatgg gtcacgacga gatcatcgcc 15240
gtcgggcatg cgcgccttga gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc 15300
ttcgtccaga tcatcctgat cgacaagacc ggcttccatc cgagtacgtg ctgctcgat 15360
gcgatgtttc gcttggtggt cgaatgggca ggtagccgga tcaagcgtat gcagccgccg 15420
cattgcatca gccatgatgg atactttctc ggcaggagca aggtgagatg acaggagatc 15480
ctgccccggc acttcgcca atagcagcca gtcccttccc gcttcagtga caacgtcgag 15540
cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg 15600
cagttcattc agggcaccgg acaggtcggg cttgacaaaa agaaccgggc gccctgcgc 15660

tgacagccgg aacacggcgg catcagagca gccgattgtc tgttggtgcc agtcatagcc 15720
gaatagcctc tccacccaag cggccggaga acctgctgc aatccatctt gttcaatcat 15780
gcgaaacgat ccagatccgg tgcagattat ttggattgag agtgaatatg agactctaata 15840
tggataccga ggggaattta tggaacgtca gtggagcatt tttgacaaga aatatttgct 15900
agctgatagt gaccttaggc gacttttgaa cgcgcaataa tggtttctga cgtatgtgct 15960
tagctcatta aactccagaa acccgcggt gagtggctcc ttcaacgttg cggttctgtc 16020
agttccaaac gtaaaacggc ttgtcccgcg tcatcgccgg gggtcataac gtgactccct 16080
taattctccg ctcatgatca gattgtcgtt tcccgccttc agtttaaact atcagtgttt 16140
gacaggatat attggcgggt aaacctaga gaaaagagcg tttattagaa taatcggata 16200
tttaaaaggc cgtgaaaagg tttatccgtt cgtccatttg tatgtgcatg ccaaccacag 16260
ggttccccag atctggcgcc ggccagcgag acgagcaaga ttggccgccc cccgaaacga 16320
tccgacagcg cggccagcac aggtgcgag gcaaattgca ccaacgcata cagcgccagc 16380
agaatgccat agtgggcggg gacgtcgtt gagtgaacca gatcgcgag gaggcccgcc 16440
agcaccggca taatcaggcc gatgccgaca gcgtcgagcg cgacagtgtc cagaattacg 16500
atcaggggta tgttgggttt cacgtctggc ctccggacca gcctccgtg gtccgattga 16560
acgcgcggat tctttatcac tgataagttg gtggacatat tatgtttatc agtgataaag 16620
tgtcaagcat gacaaagttg cagccgaata cagtgatccg tgccgccttg gacctgttga 16680
acgaggtcgg cgtagacggg ctgacgacac gcaaactggc ggaacggttg ggggttcagc 16740
agccggcgct ttactggcac ttcaggaaca agcgggcgct gctcgacgca ctggccgaag 16800
ccatgctggc ggagaatcat acgcattcgg tgccgagagc cgacgacgac tggcgctcat 16860
ttctgatcgg gaatgcccgc agcttcaggc aggcgctgct cgcctaccgc gatggcgcg 16920

gcatccatgc cggcacgcga cggggcgcac cgcagatgga aacggccgac gcgcagcttc 16980
gcttctcttg cgaggcgggt ttttcggccg gggacgccgt caatgcgctg atgacaatca 17040
gctacttcac tggtggggcc gtgcttgagg agcaggccgg cgacagcgat gccggcgagc 17100
ggggcggcac cgttgaacag gctccgctct cgccgctggt gcggggccgcg atagacgcct 17160
tcgacgaagc cgggccggac gcagcgcttc agcagggact cgcggtgatt gtcgatggat 17220
tggcgaaaag gaggctcgtt gtcaggaacg ttgaaggacc gagaaagggt gacgattgat 17280
caggaccgct gccggagcgc aaccactca ctacagcaga gccatgtaga caacatcccc 17340
tccccctttc caccgcgtca gacgcccgtg gcagcccgt acgggctttt tcatgccctg 17400
ccctagcgtc caagcctcac ggccgcgctc ggcctctctg gcggccttct ggcgctcttc 17460
cgcttctctg ctactgact cgctgcgctc ggtcgcttcgg ctgcggcgag cggtatcagc 17520
tcactcaaag gcggtaatat gggtatccac agaatcaggg gataacgcag gaaagaacat 17580
gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcggttc tggcgttttt 17640
ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc agagggtggcg 17700
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 17760
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 17820
ggcgcttttc cgctgcataa ccctgcttcg gggtcattat agcgattttt tcggtatatc 17880
catccttttt cgcacgatat acaggatttt gccaaagggt tcgtgtagac tttccttggt 17940
gtatccaacg gcgtcagccg ggcaggatag gtgaagtagg cccaccgcg agcgggtggt 18000
ccttcttcac tgctcccttat tcgcacctgg cggtgctcaa cgggaatcct gctctgcgag 18060
gctggccggc taccgccggc gtaacagatg agggcaagcg gatggctgat gaaaccaagc 18120
caaccaggaa gggcagccca cctatcaagg tgtactgcct tccagacgaa cgaagagcga 18180

ttgaggaaaa ggcggcggcg gccggcatga gcctgtcggc ctacctgctg gccgtcggcc 18240
agggtacaa aatcacgggc gtcgtggact atgagcacgt ccgcgagctg gcccgcatca 18300
atggcgacct gggccgcctg ggcggcctgc tgaaactctg gctcaccgac gacccgcgca 18360
cggcgcgggtt cggatgatgcc acgatcctcg ccctgctggc gaagatcgaa gagaagcagg 18420
acgagcttgg caaggtcatg atgggcgtgg tccgcccag ggagagcca tgactttttt 18480
agccgctaaa acggccgggg ggtgcgcgtg attgccaagc acgtcccat gcgctccatc 18540
aagaagagcg acttcgcgga gctggtgaag tacatcaccg acgagcaagg caagaccgag 18600
cgcttttgcg acgtca 18617

<210> 51

<211> 18333

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 51

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gggccgcgcg gcgttggtga tacctcgcg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt	480
tttccggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaac cctcccggcc cgctaacgcg ggctctccat cccccaggg gctgcgcccc	660
tcggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgca cgcccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac cagggtgccg gcagtgaggg cggcggcctg	840
ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140
agataatata tcttttatat agaagatac gccgtatgta aggatttcag ggggcaaggc	1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga	1260

cta atgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttctgtccatt tccgcgaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg gggtgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100
aatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccatth aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgct ttttccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgct cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatthtt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgtthtag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgttht ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggctcgtggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggtthtccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgct 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgtcttg 3840
ccctgttcac cagcgcaac aagaaaatcc cgcgcgaggc gctgcaaac aaggatcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggg gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcgggtgtgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgctcgtg ctgtttgctg 4260
gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgc 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccg gctcaagctg gaaaccttc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcatgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gtcagtatc 4620
gctcgggacg cacggcgcg cttacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccca 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccga cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattctt catcctcggc gcaacttaata tttcgctatt 5100
ctggagcttg ttgtttatatt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgt ctgctaggta gcccgatacg 5220
attgatggcg gtcctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgttgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagt gcaacctccc gtgcctctgc tcacctttac 5400
cgcctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcattcaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cggcgtccc gactgatggg ctgcctgtat 6060
cgagtgggtga ttttggtgccg agctgccgtt cggggagctg ttggctgggt ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcgggtacc cgggatcttt cgacactgaa atacgtcgag cctgctcgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900
gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctgcttgga gcatagaggg 7320
tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcggtga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaadc cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggtc agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtccctg gtaggcagct ttgccccgtc tgtccgcccc gtgtgtcggc ggggttgaca 8700

aggctcgttc gtcagtcaca catttggttc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgtctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcctcgcagc ttgactaaca gctaccccgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcctcggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtc gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgtc cggccgatct tagccagacg agcgggttcg gcccatcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggtccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgcgcacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgctgatgct agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaaat agctccatgt caacaagaat aaaacgcggtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctggggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800
tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc toatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160
cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac tctaaatctc gatgaccga ccaaaaagag cgaacaaaga aataatcttg 11280
tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400
tatgttggtg atttggaatg ccctgatcgt tttcgttacc gtgattggca tggaagtgat 11460
tgctgcactg gcacacaaat acatcatgca cggctggggg tggggatggc atctttcaca 11520
tcatgaaccg cgtaaagggt cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580
attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640
cgcaggatat acggcgatat gattactcta ttttatgggt cacgacgggc tggatgcatca 11700
acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760
ccgtatgcat cacgccgtca ggggcaaaga aggttggtgt tcttttggct tcctctatgc 11820
gccgcccctg tcaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcgggcgc 11880
tgccagagat ggcagggcg gggaggatga gcccgcatcc gggaagtaag ggcctgacca 11940
gaggcggcca gcagcagcgt taatttttcg ggcgtgggtc ttgactgccg ctgatcccaa 12000
agcttgagat taaaatagat aaggaaaaga aagtgaaaag aaattcggaa gcatggcaca 12060
ttcttctttt tataaatata tgcttgactt tctttttcca tcgatatgat atatgcatat 12120
gatagatata caagcaatct tcttcaagga gtttgaaatt ttgtcctcca ggagcaaaaa 12180
aaagtttttt ttatatactg tttgtacaca agaatagtta ccaatttgct ttggtcttac 12240
gtgctgcaag tttatatcgt tttcaatttc tttgtcttta cattttcttt gtcctttatc 12300
tttctcatt tagtcttttg gagaattagg aaaagggagc ggaaaggtaa gaaatgcttg 12360
cgtattttac taattcggca aacatccaat ttggcaaaca gcagcctgtg caacgctctc 12420
gagatgacag tatctttgat tacactctaa atctcgatga cccgaccaa aagagcgaac 12480
aaagaaataa tcttgatgcat tcgaatatga tggaagattt tttccccctt attctaaatg 12540
ttgacatagc gtgtatgtta tataaacaaa aagaaattgt acaaactttc ttttcttctc 12600

ttttttatttt atctctatga tccagttaga acaaccactc agtcatcaag caaaactgac 12660
tccagtactg agaagtaa atcagtttaa ggggcttttc attgctattg tcattgttag 12720
cgcatgggtc attagcctga gtttattact ttcccttgac atctcaaagc taaaattttg 12780
gatgttattg cctgttatac tatggcaa acattttatat acgggattat ttattacatc 12840
tcatgatgcc atgcatggcg tagtatttcc ccaaaacacc aagattaatc atttgattgg 12900
aacattgacc ctatcccttt atgggtcttt accatatcaa aaactattga aaaaacattg 12960
gttacaccac cacaatccag caagctcaat agaccgggat tttcacaatg gtaaacacca 13020
aagtttcttt gcttgggtatt ttcattttat gaaagggtac tggagttggg ggcaaataat 13080
tgcggtgact attatttata actttgctaa atacatactc catatcccaa gtgataatct 13140
aacttacttt tgggtgctac cctcgctttt aagttcatta caattattct attttggtac 13200
ttttttaccc catagtgaac caataggggg ttatgttcag cctcattgtg cccaacaat 13260
tagccgtcct atttggtggt catttatcac gtgctatcat tttggctacc acgaggaaca 13320
tcacgaatat cctcatattt cttggtggca gttaccagaa atttacaag caaaatagaa 13380
gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc 13440
cacacaacat acgagccgga agcataaagt gtaaagcctg ggtgcctaa tgagttagct 13500
aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc 13560
agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggccaaaga 13620
caaaagggcg acattcaacc gattgagga ggaaggtaa atattgacgg aaattattca 13680
ttaaagggtga attatcacg tcaccgactt gagccatttg ggaattagag ccagcaaat 13740
caccagtagc accattacca ttagcaaggc cggaacgct accaatgaaa ccatcgatag 13800
cagcaccgta atcagtagcg acagaatcaa gtttgccttt agcgtcagac ttagcgcgt 13860

tttcatcggc attttcggtc atagccccct tattagcggt tgccatcttt tcataatcaa 13920
aatcaccgga accagagcca ccaccggaac cgctccctc agagccgcca cctcagaac 13980
cgccaccctc agagccacca cctcagagc cgccaccaga accaccacca gagccgccgc 14040
cagcattgac aggaggcccg atctagtaac atagatgaca ccgcgcgcga taatttatcc 14100
tagtttgccg gctatatttt gttttctatc gcgtattaaa tgtataattg cgggactcta 14160
atcataaaaa cccatctcat aaataacgtc atgcattaca tgttaattat tacatgctta 14220
acgtaattca acagaaatta tatgataatc atcgcaagac cggcaacagg attcaatctt 14280
aagaaacttt attgccaaat gtttgaacga tcggggatca tccgggtctg tggcgggaac 14340
tccacgaaaa tatccgaacg cagcaagata tcgcggtgca tctcggtctt gcctgggcag 14400
tcgccgccga cgccgttgat gtggacgccg ggcccgatca tattgtcgct caggatcggt 14460
gcgttggtgt tgtcggccgt tgctgtcgta atgatatcgg caccttcgac cgctgttcc 14520
gcagagatcc cgtgggcgaa gaactccagc atgagatccc cgcgctggag gatcatccag 14580
ccggcgctccc ggaaaacgat tccgaagccc aacctttcat agaaggcggc ggtggaatcg 14640
aaatctcgtg atggcaggtt gggcgctcgt tggtcggtca tttcgaacct cagagtcccg 14700
ctcagaagaa ctcgtaaga aggcgataga aggcgatgcg ctgcgaatcg ggagcggcga 14760
taccgtaaag cagaggaag cggtcagccc attcgccgcc aagctcttca gcaatatcac 14820
gggtagccaa cgctatgtcc tgatagcggc ccgccacacc cagccggcca cagtcgatga 14880
atccagaaaa ggggccattt tccaccatga tattcggcaa gcaggcatcg ccatgggtca 14940
cgacgagatc atcgccgtcg ggcgatgcgc ccttgagcct ggcgaaacagt tcggctggcg 15000
cgagccctg atgctcttcg tccagatcat cctgatcgac aagaccggct tccatccgag 15060
tacgtgctcg ctcgatgcga tgtttcgctt ggtggctgaa tgggcaggta gccggatcaa 15120

gcgtatgcag cgcgcgcatt gcatcagcca tgatggatac tttctcggca ggagcaaggt 15180
gagatgacag gagatcctgc cccggcactt cgccaatag cagccagtcc cttcccgtt 15240
cagtgacaac gtcgagcaca gctgcgcaag gaacgcccgt cgtggccagc cacgatagcc 15300
gcgctgcctc gtccctgcagt tcattcaggg caccggacag gtcggtcttg acaaaaagaa 15360
ccgggcgccc ctgcgctgac agccggaaca cggcggcatc agagcagccg attgtctgtt 15420
gtgcccagtc atagccgaat agcctctcca cccaagcggc cggagaacct gcgtgcaatc 15480
catcttggtc aatcatgcga aacgatccag atccggtgca gattatttgg attgagagtg 15540
aatatgagac tctaattgga taccgagggg aatttatgga acgtcagtgg agcatttttg 15600
acaagaaata tttgctagct gatagtgacc ttaggcgact tttgaacgcg caataatggt 15660
ttctgacgta tgtgcttagc tcattaaact ccagaaacct gcggctgagt ggctccttca 15720
acgttgcggt tctgtcagtt ccaaacgtaa aacggcttgt cccgcgtcat cggcgggggt 15780
cataacgtga ctcccttaat tctccgtca tgatcagatt gtcgtttccc gccttcagtt 15840
taaactatca gtgtttgaca ggatatattg gcgggtaaac ctaagagaaa agagcgttta 15900
ttagaataat cggatattta aaagggcgtg aaaagggtta tccgttcgtc catttgatg 15960
tgcatgccaa ccacagggtt cccagatctt ggcgccggcc agcgagacga gcaagattgg 16020
ccgccgcccc aaacgatccg acagcgcgcc cagcacaggt gcgcaggcaa attgcaccaa 16080
cgcatacagc gccagcagaa tgccatagtg ggcggtgacg tcgttcgagt gaaccagatc 16140
gcgcaggagg cccggcagca ccggcataat caggccgatg ccgacagcgt cgagcgcgac 16200
agtgtcaga attacgatca ggggtatggt gggtttcacg tctggcctcc ggaccagcct 16260
ccgctgggtc gattgaacgc gcggattctt tatcactgat aagttggtgg acatattatg 16320
tttatcagtg ataaagtgtc aagcatgaca aagttgcagc cgaatacagt gatccgtgcc 16380

gccctggacc tgttgaacga ggtcggcgta gacggtctga cgacacgcaa actggcgga 16440
cggttggggg ttcagcagcc ggcgctttac tggcacttca ggaacaagcg ggcgctgctc 16500
gacgcactgg ccgaagccat gctggcggag aatcatacgc attcggtgcc gagagccgac 16560
gacgactggc gctcatttct gatcggaat gcccgagct tcaggcaggc gctgctcgcc 16620
taccgcatg gcgcgcgcat ccatgccggc acgcgacgg gcgcaccgca gatggaaacg 16680
gccgacgcgc agcttcgctt cctctgcgag gcgggttttt cggccgggga cgccgtcaat 16740
gcgctgatga caatcagcta cttcactgtt ggggccgtgc ttgaggagca ggccggcgac 16800
agcgatgccg gcgagcgcg gcgcaccgtt gaacaggctc cgctctcgcc gctgttgcg 16860
gccgcatag acgccttcga cgaagccggt ccggacgcag cgctcgagca gggactcgcg 16920
gtgattgtcg atggattggc gaaaaggagg ctggtgtca ggaacgttga aggaccgaga 16980
aagggtgacg attgatcagg accgctgccg gagcgcaacc cactcactac agcagagcca 17040
tgtagacaac atccccccc cctttccacc gcgtcagacg ccgtagcag cccgctacgg 17100
gctttttcat gccctgccct agcgtccaag cctcacggcc gcgctcgcc tctctggcg 17160
ccttctggcg ctcttcgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 17220
ggcgagcgg atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 17280
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 17340
cggttgctggc gtttttccat aggctccgcc .cccctgacga gcatcacaaa aatcgacgct 17400
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt cccctggaa 17460
gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgccttcc 17520
tcccttcggg aagcgtggcg cttttccgct gcataaccct gcttcgggggt cattatagcg 17580
attttttcgg tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt 17640

gtagactttc cttggtgtat ccaacggcgt cagccgggca ggataggtga agtaggccca 17700
cccgcgagcg ggtgttcctt cttcactgtc ccttattcgc acctggcggt gctcaacggg 17760
aatcctgctc tgcgaggctg gccggctacc gccggcgtaa cagatgaggg caagcggatg 17820
gctgatgaaa ccaagccaac caggaagggc agcccaccta tcaaggtgta ctgccttcca 17880
gacgaacgaa gagcgattga ggaaaaggcg gcggcgcccg gcatgagcct gtcggcctac 17940
ctgctggccg tcggccaggg ctacaaaatc acgggcgtcg tggactatga gcacgtccgc 18000
gagctggccc gcatcaatgg cgacctgggc cgcctgggcg gcctgctgaa actctggctc 18060
accgacgacc cgcgcacggc gcggttcggt gatgccacga tcctcgccct gctggcgaag 18120
atcgaagaga agcaggacga gcttggcaag gtcattgatg gcgtgggtccg cccgagggca 18180
gagccatgac ttttttagcc gctaaaacgg ccgggggggtg cgcgtgattg ccaagcacgt 18240
ccccatgcgc tccatcaaga agagcgactt cgcgagctg gtgaagtaca tcaccgacga 18300
gcaaggcaag accgagcgcc tttgcgacgc tca 18333

<210> 52

<211> 17

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (9)..(9)

<223> n is a, c, g, or t

<400> 52

gcngarggna thtggta

17

<210> 53

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (6)..(6)

<223> n is a, c, g, or t

<400> 53

tcngcnagra adatrtrtrtg

20

<210> 54

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 54

aagtgacacc gggtacacgc ttgtctt

27

<210> 55

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 55

gcttatcacc atctggtacc tccttgc

27

<210> 56

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 56

agagagggat ccttaaagtc gaatatcggt gc

32

<210> 57

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 57

agagagggat ccatgtctga tcaaaagaag ca

32

<210> 58

<211> 37

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 58

actttattgg atccttaa at gcgaatatcg ttgctgc

37

<210> 59

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 59

gttccaattg gccacatgaa gagtaagaca ggaaacag

38

<210> 60

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 60

cctgtcttac tcttcatgtg gccaatgga accaacac

38

<210> 61

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 61

ctatttta at catatgtctg atcaaaagaa gcatattg

38

<210> 62
<211> 16103
<212> DNA
<213> Artificial

<220>
<223> Primer

<220>
<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 62
gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60

cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120

cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180

cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240

tcgggaagcc agcgactggg gacctccagc gtccctttaa ggctgccaac agctttctca 300

gccagggcca gccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360

gtgggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg 420

atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

tttggaattg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc 540

aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600

gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660

ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720

tccgatagct ctgcaaaggc cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780

tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagccctt acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgacct agaggggtcat gacttgagcc taaaatccgc cgctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttcttg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgagggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacggaga atatggagct 1500

tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttcgatc tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgct caaatatcgt gcctctcctg ctttgcccgg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860
ggcagctttg ccccgctctgt ccgcccgggtg tgtcggcggtg gttgacaagg tcgttgcgctc 1920
agtccaacat ttgttgccat attttccctgc tctccccacc agctgctctt ttctttttctc 1980
tttctttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040
taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100
gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
accatgcctg aactcacgcg gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctcgaggggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcggtg gatatgtcct gcgggtaa at agctgcgcgc atggttttcta caaagatcgt 2340
tatgtttatc ggcactttgc atcgggccgc ctcccgattc cggaagtgtc tgacattggg 2400
gaattcagcg agagcctgac ctattgcac tcccgccgtg cacaggggtgt cacgttgcaa 2460
gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580
gggtcaatata ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
tggcaaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga ttctggctcc 2760
aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820
ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcacccg agcttgcagg atcgccgcgg 2940
ctccgggcgt atatgctccg cattggtctt gaccaactct atcagagctt ggttgacggc 3000
aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctccg atccggagcc 3060

gggactgtcg ggcgtacaca aatcgcccg cagaagcgcgg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgccccagca ctcggtccgag ggcaaaggaa 3180

tagagtagat gccgaccg cggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tggttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ctttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540

tcaacgggtcg tcaagagacc tacgagactg aggaatccgc tcttggctcc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780

gatggtatac aatagtaacc atgcatgggt gcttagtgaa tgctccgtaa cacccaatac 3840

gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900

cttggttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct ccaactcgacc tgcaggcatg caagcttgag tctatgcct ccaaaaagta 4020

cggtgctgaa ttcagatata aatcgctgtg tgctaaaatt aacactgtcg ataaagacaa 4080

gcgtgtaacc ggtgtcactt tggaaagcgg agaagtcatt gaagccgatg cagtcgtatg 4140

taatgcggat cttgtttatg cttatcacca tctgttacct ctttgcaatt ggacaaagaa 4200

gacattagcc tcaaagaaac tcaacttcac atctatttcg ttttattgggt ccatgtcaac 4260

aaaggtgcct caattagacg tacacaatat cttcttggct gaagcctaca aggaaagttt 4320

tgatgagatt ttcaacgact tcggtttgcc ctctgaagct tggcgtaatc atgggtcatag 4380

ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc 4440

ataaagtgta aagcctgggg tgcctaata gaagagtaac tcacattaat tgcgttgccg 4500

tcactgcccc ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa 4560

cgcgcggggga gaggcgggtt gcgtattggg ccaaagacaa aagggcgaca ttcaaccgat 4620

tgagggaggg aaggtaaata ttgacggaaa ttattcatta aaggtgaatt atcacgtca 4680

ccgacttgag ccatttgagg attagagcca gcaaaatcac cagtagcacc attaccatta 4740

gcaaggccgg aaacgtcacc aatgaaacca tcgatagcag caccgtaatc agtagcgaca 4800

gaatcaagtt tgccttttagc gtcagactgt agcgcgtttt catcggcatt ttcggtcata 4860

gcccccttat tagcgtttgc catcttttca taatcaaat caccggaacc agagccacca 4920

ccggaaccgc ctccctcaga gccgccacc tcagaaccgc caccctcaga gccaccacc 4980

tcagagccgc caccagaacc accaccagag ccgccgccag cattgacagg aggcccgatc 5040

tagtaacata gatgacaccg cgcgcgataa tttatcctag ttgcgcgct atattttgtt 5100

ttctatcgcg tattaaatgt ataattgggg gactctaata ataaaaacc attcataaa 5160

taacgtcatg cattacatgt taattattac atgcttaacg taattcaaca gaaattatat 5220

gataatcatc gcaagaccgg caacaggatt caatcttaag aaactttatt gccaaatgtt 5280

tgaacgatcg gggatcatcc ggggtctgtg cggaactcc acgaaaatat ccgaacgcag 5340

caagatatcg cggtgcatct cggtcttgcc tgggcagtcg ccgccgacgc cgttgatgtg 5400

gacgccgggc ccgatcatat tgtcgctcag gatcgtggcg ttgtgcttgt cggccgttgc 5460

tgtcgtaatg atatcggcac cttcgaccgc ctgttccgca gagatcccgt gggcgaagaa 5520

ctccagcatg agatccccgc gctggaggat catccagccg gcgtcccgga aaacgattcc 5580

gaagcccaac ctttcataga aggcggcggt ggaatcgaaa tctcgtgatg gcaggttggg 5640
cgtcgcttgg tcggtcattt cgaaccccag agtcccgtc agaagaactc gtcaagaagg 5700
cgatagaagg cgatgcgctg cgaatcgga ggcggcgatac cgtaaagcac gaggaagcgg 5760
tcagcccat cgcgcgaag ctcttcagca atatcacggg tagccaacgc tatgtcctga 5820
tagcgggtccg ccacaccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc 5880
accatgatat tcggcaagca ggcacgcga tgggtcacga cgagatcatc gccgtcgggc 5940
atgcgcgcct tgagcctggc gaacagttcg gctggcgca gccctgatg ctcttcgtcc 6000
agatcatcct gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt 6060
ttcgcttggg ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca 6120
tcagccatga tggatacttt ctccgcagga gcaaggtgag atgacaggag atcctgcccc 6180
ggcacttcgc ccaatagcag ccagtcctt cccgcttcag tgacaacgct gagcacagct 6240
gcgcaaggaa cgcgcgtcgt ggccagccac gatagccgcg ctgcctcgct ctgcagttca 6300
ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc 6360
cggaacacgg cggcatcaga gcagccgatt gtctgttgtg ccagtcata gccgaatagc 6420
ctctccacc aagcggccgg agaacctgcg tgcaatccat cttgttcaat catgcgaaac 6480
gatccagatc cgggtgcagat tatttggatt gagagtgaat atgagactct aattggatac 6540
cgaggggaat ttatggaacg tcagtggagc atttttgaca agaaatattt gctagctgat 6600
agtgcctta ggcgactttt gaacgcgcaa taatggtttc tgacgtatgt gcttagctca 6660
ttaaactcca gaaacccgcg gctgagtggc tccttcaacg ttgcggttct gtcagttcca 6720
aacgtaaac ggcttgtccc gcgtcatcgg cgggggtcat aacgtgactc ccttaattct 6780
ccgctcatga tcagattgtc gtttcccgcc ttcagtttaa actatcagtg tttgacagga 6840

tatattggcg ggtaaacct agagaaaaga gcgtttatta gaataatcgg atatttaaaa 6900

gggcgtgaaa aggtttatcc gttcgtccat ttgtatgtgc atgccaacca cagggttccc 6960

cagatctggc gccggccagc gagacgagca agattggccg ccgcccga aa cgatccgaca 7020

gcgcgcccag cacaggtgcg caggcaaatt gcaccaacgc atacagcgcc agcagaatgc 7080

catagtgggc ggtgacgtcg ttcgagtga ccagatcgcg caggaggccc ggcagcaccg 7140

gcataatcag gccgatgccg acagcgtcga gcgcgacagt gtcagaatt acgatcaggg 7200

gtatgttggg ttacacgtct ggcctccgga ccagcctccg ctggtccgat tgaacgcgcg 7260

gattctttat cactgataag ttggtggaca tattatgttt atcagtgata aagtgtcaag 7320

catgacaaag ttgcagccga atacagtgat ccgtgccgcc ctggacctgt tgaacgaggt 7380

cggcgtagac ggtctgacga cagcgaact ggcggaacgg ttgggggttc agcagccggc 7440

gctttactgg cacttcagga acaagcgggc gctgctcgac gcactggccg aagccatgct 7500

ggcggagaat catacgcat cgggtccgag agccgacgac gactggcgct catttctgat 7560

cgggaatgcc cgcagcttca ggcaggcgct gctcgctac cgcgatggcg cgcgcatcca 7620

tgccggcacg cgaccgggcg caccgcagat ggaaacggcc gacgcgcagc ttcgcttcc 7680

ctgcgaggcg ggtttttcgg ccggggacgc cgtcaatgcg ctgatgacaa tcagctactt 7740

cactgttggg gccgtgcttg aggagcaggc cggcgacagc gatgccggcg agcgcggcgg 7800

caccgttgaa caggctccgc tctcgccgct gttgcgggccc gcgatagacg ccttcgacga 7860

agccgggtccg gacgcagcgt tcgagcaggg actcgcggtg attgtcgatg gattggcgaa 7920

aaggaggctc gttgtcagga acgttgaagg accgagaaag ggtgacgatt gatcaggacc 7980

gctgcgggag cgcaaccac tcactacagc agagccatgt agacaacatc ccctccccct 8040

ttccaccgcg tcagacgccc gtagcagccc gctacgggct ttttcatgcc ctgccctagc 8100

gtccaagcct cacggccgcg ctcggcctct ctggcggcct tctggcgctc ttccgcttcc 8160

tcgctcactg actcgctgcg ctcggtcggt cggctgcggc gagcggtatc agctcactca 8220

aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 8280

aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg 8340

ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 8400

acaggactat aaagatacca ggcgtttccc cctggaagct cctcgtgcg ctctcctggt 8460

ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 8520

ttccgctgca taacctgct tcgggggtcat tatagcgatt ttttcggtat atccatcctt 8580

tttcgcacga tatacaggat ttgccaaag ggttcgtgta gactttcctt ggtgtatcca 8640

acggcgctcag cggggcagga taggtgaagt agggccaccc gcgagcgggt gttccttctt 8700

cactgtccct tattcgcacc tggcggtgct caacgggaat cctgctctgc gaggctggcc 8760

ggctaccgcc ggcgtaacag atgagggcaa gcggatggct gatgaaacca agccaaccag 8820

gaagggcagc ccacctatca aggtgtactg ctttcagac gaacgaagag cgattgagga 8880

aaaggcggcg gcggccggca tgagcctgtc ggcctacctg ctggccgctc gccagggcta 8940

caaaatcacg ggcgtcgtgg actatgagca cgtccgcgag ctggcccga tcaatggcga 9000

cctgggccgc ctgggcggcc tgctgaaact ctggctcacc gacgaccgc gcacggcgcg 9060

gttcggtgat gccacgatcc tcgccctgct ggcgaagatc gaagagaagc aggacgagct 9120

tggcaaggtc atgatggcg tggtcgccc gagggcagag ccatgacttt tttagccgct 9180

aaaacggccg gggggtgcgc gtgattgcca agcacgtccc catgcgctcc atcaagaaga 9240

gcgacttcgc ggagctggtg aagtacatca ccgacgagca aggcaagacc gagcgccttt 9300

gcgacgctca ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca 9360

aacgcgccag aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga 9420

tacctcgcg aaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg 9480

gccgactcac ccggcgcggtc gttgacagat gaggggcagg ctcgatttcg gccggcgacg 9540

tggagctggc cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag 9600

atgatgtgga caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact 9660

actgacagat gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg 9720

gcgcacctat tgacatttga ggggtgtcc acaggcagaa aatccagcat ttgcaagggt 9780

ttccgcccgt ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt 9840

ataaaccttg tttttaacca gggctgcgcc ctgtgcgctg gaccgcgcac gccgaagggg 9900

ggtgcccccc cttctcgaa cctcccgcc cgctaacgcg ggcctcccat cccccaggg 9960

gctgcgcccc tcggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc 10020

cattgccggg atcggggcag taacgggatg ggcgatcagc ccgagcgga cggccggaag 10080

cattgacgtg ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg 10140

cggcggcctg ggtggcggtc tgcccttcac ttcgccgctc ggggcattca cggacttcat 10200

ggcggggccg gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgcccgtgt 10260

cgtgttcggg ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataaccg 10320

aggtatgaaa acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa 10380

agctaccaag acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac 10440

aatactgata agataatata tcttttatat agaagatatc gccgtatgta aggatttcag 10500

ggggcaaggc ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa 10560

cttgcattga ctaatgcttg aaaccagga caataacctt atagcttgta aattctatca 10620

taattgggta atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac 10680
tttgtcatgc agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag 10740
gtgctgcctc agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac 10800
gtgcagcttt cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac 10860
cacgtcaaag ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc 10920
gaatacgtgc gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg 10980
gcgcgattta gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc 11040
actgcccggc tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa 11100
atcgtgttga ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg 11160
gccatatcaa tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt 11220
tgccatgttt tacggcagtg agagcagaga tagcgtgat gtccggcggt gcttttgccg 11280
ttacgcacca ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact 11340
ggagcacctc aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat 11400
tgtggtttca aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt 11460
gaaaaagctg ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc 11520
gtcttgttat aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat 11580
aataaatggc taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct 11640
gcgtaaaaga tacggaagga atgtctcctg ctaagggtata taagctggtg ggagaaaatg 11700
aaaacctata tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac 11760
gggaaaagga catgatgcta tggctggaag gaaagctgcc tgttccaaag gtctgcact 11820
ttgaacggca tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct 11880

cggaagagta tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc 11940
tcaggctctt tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc 12000
gcttagccga attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact 12060
gggaagaaga cactccattt aaagatccgc gcgagctgta tgatTTTTTA aagacggaaa 12120
agcccgaaga ggaacttgtc TTTTcccacg gcgacctggg agacagcaac atctttgtga 12180
aagatggcaa agtaagtggc ttattgatc ttgggagaag cggcagggcg gacaagtgg 12240
atgacattgc cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg 12300
agctatTTTT tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt 12360
tactggatga attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag 12420
cgcaccgact tcttccgcat caagtgtttt ggctctcagg ccgaggcca cggcaagtat 12480
ttgggcaagg ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag 12540
gacggccaga cggctacgg gaccgacttc attgccgata aggtggatta tctggacacc 12600
aaggcaccag gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca 12660
atcccgcaag gagggatgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg 12720
atcgacgcgg ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgct 12780
gcgccccgcg aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc 12840
gagcgcgaca gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag 12900
cgttcgcgtc gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg 12960
cgaggaacta tgacgaccaa gaagcga aaa accgccggcg aggacctggc aaaacaggtc 13020
agcgaggcca agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag 13080
ctttccttgt tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg 13140

gcccgtcttg ccctgttcac cagcgcgaac aagaaaatcc cgcgcgaggc gctgcaaaac 13200
aaggtcattt tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg 13260
gccgacgatg acgaactggg gtggcagcag gtgttgaggc acgcgaagcg caccctatc 13320
ggcgagccga tcaccttcac gttctacgag ctttgccagg acctgggctg gtogatcaat 13380
ggccggtatt acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc 13440
ttcacgtccg accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc 13500
ctggaccgtg gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg 13560
ctgtttgctg gcgaccacta caggaattc atatgggaga agtaccgcaa gctgtcgccg 13620
acggcccgac ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg 13680
gaaaccttcc gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag 13740
gtcggcgaag cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat 13800
gatgacctgg tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca 13860
gcagccagcg ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc 13920
gctcagtatc gctcgggacg cagggcgcgc tctacgaact gccgataaac agaggattaa 13980
aattgacaat tgtattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt 14040
tccgcgagat ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg 14100
agcacgagga gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat 14160
tcggcgcccta catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc 14220
ccaaggacgc tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc 14280
gaggggtcgc cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg 14340
tccgacagat tccaacggga atctggtgga tgcgcattc catcctcggc gcacttaata 14400

tttcgctatt ctggagcttg ttgtttatth cggctctaccg cctgccgggc ggggtcgcgg 14460
cgacggtagg cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta 14520
gcccatacag attgatggcg gtcctggggg ctatttgccg aactgccggc gtggcgctgt 14580
tggtgttgac accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg 14640
cggtttccat ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc 14700
tcacctttac cgctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag 14760
tgtttgatcc gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg 14820
gcctgatcgg agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac 14880
ctacagttgt ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga 14940
tgcatcaggc cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg 15000
ataggggagt tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc 15060
agcggcttta tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt 15120
cacggttaag cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga 15180
tatttgatca caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca 15240
gatcatccgt gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac 15300
atgagcaaag tctgccgcct tacaacggct ctcccgtga cgcgcgccg gactgatggg 15360
ctgcctgtat cgagtgggta ttttgtgccg agctgccggg cggggagctg ttggctggct 15420
gggtggcagga tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg 15480
cggacgtttt taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg 15540
attgcccttc accgcctggc cctgagagag ttgcagcaag cgggtccaogc tggtttgccc 15600
cagcaggcga aaatcctgtt tgatgggtgg tccgaaatcg gcaaaatccc ttataaatca 15660

aaagaatagc ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta 15720
aagaacgtgg actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta 15780
cgtgaaccat cacccaaatac aagttttttg gggtcgaggt gccgtaaagc actaaatcgg 15840
aaccctaaag ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga 15900
aaggaagga agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg 15960
gcgatcggtg cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag 16020
gcgattaagt tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag 16080
tgaattcgag ctcggtaccc ggg 16103

<210> 63

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 63

ggcgtacttg aaggaaccct taccg

25

<210> 64

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 64

attgatgctc ccggtcaccg tgatt

25

<210> 65

<211> 500

<212> DNA

<213> *Blakeslea trispora*

<400> 65

aatctataca atgctccata gactcacatt gatattgtcg aagatttcga tgctgactta 60
gtagagcaac tacaaaagtt agcagagaag catgatttct taatctttga agaccgcaag 120
tttgagata tcggtatgtg aattctatct attttttttc tgatgtgtgc atggatgact 180
catgatcata ttcttaggta atactgtcaa gcatcaatat ggcaagggcg ttacaagat 240
tgcttcttgg tctcatatta ctaatgctca cacagttcct ggagaaggta ttatcaaggg 300
acttgccgaa gtcggcctcc ctcttggtcg tggcttgctt ttgctagcag aaatgtcatc 360
tcaagggtgca ttaactaagg gtatttacac tgccgaatct gtcaatatgg ctgcccgcaa 420
caaagatttc gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga 480
ttttgttgtc atgtcgctg 500

<210> 66

<211> 611

<212> DNA

<213> *Blakeslea trispora*

<400> 66

gagattaaaa tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt 60
ctttttataa atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag 120
atatacaagc aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt 180
ttttttttat acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct 240
gcaagtttat atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc 300
tcatttagtc tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat 360

tttactaatt cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat 420
gacagtatct ttgattacac tctaaatctc gatgaccgga ccaaaaagag cgaacaaaga 480
aataatcttg tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac 540
atagcgtgta tggtatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt 600
attttatctc t 611

<210> 67

<211> 720

<212> DNA

<213> Blakeslea trispora

<400> 67

atgtcaatac tcacttatct ggaatttcat ctctactata cactacctgt ccttgcgcca 60
ttgtgttggc tgctaaagcc gtttcactca cagcaagaca atctcaagta taaattttta 120
atgttgatgg ccgcctctac cgcacgatt tgggacaatt atatcgttta tcatcgcgct 180
tggtgggtact gtctacttg tgttgtggct gtcattggct atgtacctct agaagaatac 240
atgttcttta tcatcatgac tttaatgact gtcgcgttct caaactttgt tatgcgttgg 300
cacttgcata ctttctttat tagaccaac acttcttggga agcaaact attagtacgc 360
cttgtgcctg tttcagcttt attggcaatc acttatcatg cttggcactt gacactgcca 420
aataaacctt cattttatgg ttcatgcac ctttgggatg cttgtcctgt gttggctatt 480
ctttggctgg gtgctggcga atatatcttg cgtcgacctg tggctgtcct tttgtctatt 540
gttatcccta gtgtatacct atgttgggct gatatcgctg ctattagtgc tggcacatgg 600
catatttctc ttagaacaag cactggcaaa atggtagtac ccgatttacc tgtagaagaa 660
tgctgtttt ttactttgat caacacagtc ttggtttttg ctacctgtgc tatagaccgc 720

<210> 68

<211> 1089

<212> DNA

<213> *Blakeslea trispora*

<400> 68

ctgtacaaat catctgttca aaatcaaaac cctaaacaag ccatttccct tttccagcat	60
gtcaaagagc tagcatgggc cttctgtctt cctgaccaa tgctcaaca tgaattgttt	120
gatgatctta ctatcagctg ggatatttta cgtaaagcct caaagtcatt ctatactgca	180
tctgccgttt ttccaagtta tgtacgtcaa gacttgggtg ttctctatgc tttctgcaga	240
gctaccgatg acctgtgcga tgatgaatcc aaatctgttc aagaaagaag agaccaatta	300
gatcttactc gacaatttgt tcgtgatctc tttagccaaa agaccagtgc gcctattgtg	360
attgattggg aattgtatca aaaccaactt cctgcttctt gtatatcagc ctttagagcc	420
tttactcgcc ttcgccatgt ccttgaagta gaccctgtag aagaactatt agatggttac	480
aaatgggatc ttgagcgtcg tcctatcctt gatgaacaag acttggaggc atactctgct	540
tgtgtggcca gtagtgtggg tgaaatgtgc acacgtgtga ttcttgctca agaccaaag	600
gaaaatgatg cttggataat tgaccgtgca cgtgagatgg ggctgggtgct acaatacgtt	660
aacattgctc gagacattgt gactgatagc gagactctgg gtcgatgtta tctgcctcaa	720
caatggctta gaaaagaaga aacagaacaa atacagcaag gcaacgccc tagcctaggt	780
gatcaaagac tgttgggctt gtctctgaag cttgtaggaa aggcagacgc tatcatgggtg	840
agagctaaga agggcattga caagttgccg gcaaactgtc aaggcggtgt acgagctgct	900
tgccaagtat atgctgcaat tggatctgta ctcaagcagc agaagacaac atatcctaca	960
agagctcatc taaaaggaag cgaacgtgcc aagattgctc tgttgagtgt atacaacctc	1020
tatcaatctg aagacaagcc tgtggctctc cgtcaagcta gaaagattaa gagttttttt	1080

gttgattag

1089

<210> 69

<211> 611

<212> DNA

<213> *Blakeslea trispora*

<400> 69

agagataaaa taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac 60

atacacgcta tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc 120

acaagattat ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca 180

aagatactgt catctcgaga gcgttgacaca ggctgctggt tgccaaattg gatgtttgcc 240

gaattagtaa aatacgcaag cattttcttac ctttccgctc ccttttccta attctcccaa 300

agactaaatg aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga 360

tataaacttg cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaacatg 420

tataaaaaaa aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat 480

tgcttgatata tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta 540

tttataaaaa gaagaatgtg ccatgcttcc gaattttctt tcactttctt ttccttatct 600

attttaatct c 611

<210> 70

<211> 882

<212> DNA

<213> *Haematococcus pluvialis*

<400> 70

atgctgtcga agctgcagtc aatcagcgtc aaggcccgcc gcgttgaact agcccgcgac 60

atcacgcggc ccaaagtctg cctgcatgct cagcgggtgct cgtagttcg gctgcgagtg 120

gcagcaccac agacagagga ggcgctggga accgtgcagg ctgccggcgc gggcgatgag 180
cacagcgccg atgtagcact ccagcagctt gaccgggcta tcgcagagcg tcgtgcccgg 240
cgcaaacggg agcagctgtc ataccaggct gccgccattg cagcatcaat tggcgtgtca 300
ggcattgcc a tcttcgccac ctacctgaga tttgccatgc acatgaccgt gggcggcgca 360
gtgccatggg gtgaagtggc tggcactctc ctcttggtgg ttggtggcgc gctcggcatg 420
gagatgtatg cccgctatgc acacaaagcc atctggcatg agtcgcctct gggctggctg 480
ctgcacaaga gccaccacac acctcgact ggaccctttg aagccaacga cttgtttgca 540
atcatcaatg gactgcccgc catgctcctg tgtacctttg gcttctggct gcccaacgtc 600
ctggggggcgg cctgctttgg agcggggctg ggcatacgc tatacggcat ggcatatatg 660
tttgtacag atggcctggt gcacaggcgc tttcccaccg ggcccatcgc tggcctgccc 720
tacatgaagc gcctgacagt ggcccaccag ctacaccaca gcggcaagta cgggtggcgcg 780
ccctggggta tgttcttggg tccacaggag ctgcagcaca ttccagggtgc ggcggaggag 840
gtggagcgac tggtcctgga actggactgg tccaagcggg ag 882

<210> 71

<211> 528

<212> DNA

<213> Erwinia uredovora

<400> 71

atgttgtgga tttggaatgc cctgatcgtt ttcgttaccg tgattggcat ggaagtgatt 60
gctgcactgg cacacaaata catcatgcac ggctgggggtt ggggatggca tctttcacat 120
catgaaccgc gtaaagggtgc gtttgaagtt aacgatcttt atgccgtggg ttttgctgca 180
ttatcgatcc tgctgattta tctgggcagt acaggaatgt ggccgctcca gtggattggc 240

343/357

gcaggatatga cggcgtatgg attactctat tttatggtgc acgacgggct ggtgcatcaa 300
cgttggccat tccgctatat tccacgcaag ggctacctca aacggttgta tatggcgcac 360
cgtatgcac acgccgtcag gggcaaagaa ggttggtgtt cttttggctt cctctatgcg 420
ccgcccctgt caaaacttca ggcgacgctc cgggaaagac atggcgctag agcgggcgct 480
gccagagatg cgcagggcgg ggaggatgag cccgcatccg ggaagtaa 528

<210> 72

<211> 762

<212> DNA

<213> Nostoc sp. PCC73102

<400> 72

atgatccagt tagaacaacc actcagtcac caagcaaaac tgactccagt actgagaagt 60
aaatctcagt ttaaggggct tttcattgct attgtcattg ttagcgcatg ggtcattagc 120
ctgagtttat tactttccct tgacatctca aagctaaaat tttggatggt attgcctggt 180
atactatggc aaacatTTTT atatacggga ttatttatta catctcatga tgccatgcat 240
ggcgtagtat ttccccaaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 300
ctttatgggc ttttaccata tcaaaaaacta ttgaaaaaac attgggttaca ccaccacaat 360
ccagcaagct caatagaccc ggattttcac aatggtaaac accaaagttt ctttgcttgg 420
tattttcatt ttatgaaagg ttactggagt tgggggcaaa taattgcgtt gactattatt 480
tataactttg ctaaatacat actccatctc ccaagtgata atctaactta cttttgggtg 540
ctaccctcgc ttttaagttc attacaatta ttctattttg gtactttttt acccatagct 600
gaaccaatag ggggttatgt tcagcctcat tgtgccc aaa caattagccg tcctatttgg 660
tggtcattta tcacgtgcta tcattttggc taccacgagg aacatcacga atatcctcat 720
atttcttggg ggcagttacc agaaatttac aaagcaaaat ga 762

<210> 73

<211> 617

<212> DNA

<213> *Haematococcus pluvialis*

<400> 73

taggggtgcgg aaccaggcac gctgggtttca cacctcatgc ctgtgataag gtgtggctag 60
agcgatgcgt gtgagacggg tatgtcacgg tcgactggtc tgatggccaa tggcatcggc 120
catgtctggt catcacgggc tgggtgcctg ggtgaagggtg atgcacatca tcatgtgcgg 180
ttggagggggc tggcacagtg tgggctgaac tggagcagtt gtccaggctg gcgttgaatc 240
agtgagggtt tgtgattggc ggttgtgaag caatgactcc gcccatattc tatttgtggg 300
agctgagatg atggcatgct tgggatgtgc atggatcatg gtagtgcagc aaactatatt 360
cacctagggc tgttggtagg atcaggtgag gccttgcaca ttgcatgatg tactcgtcac 420
ggtgtgttgg tgagaggatg gatgtggatg gatgtgtatt ctgagacgta gaccttgact 480
ggaggcttga tcgagagagt gggccgtatt ctttgagagg ggaggctcgt gccagaaatg 540
gtgagtggat gactgtgacg ctgtacattg caggcagggtg agatgcactg tctcgattgt 600
aaaatacatt cagatgc 617

<210> 74

<211> 1208

<212> DNA

<213> *Haematococcus pluvialis*

<400> 74

attgtgactg atagcgagac tctgggtcga tggtatctgc ctcaacaatg gcttagaaaa 60
gaagaaacag aacaaataca gcaaggcaac gcccgtagcc taggtgatca aagactgttg 120
ggcttgtctc tgaagcttgt aggaaaggca gacgctatca tgggtgagagc taagaagggc 180

attgacaagt tgccggcaaa ctgtcaaggc ggtgtacgag ctgcttgcca agtatatgct 240
gcaattggat ctgtactcaa gcagcagaag acaacatata ctacaagagc tcatctaaaa 300
ggaagcgaac gtgccaagat tgctctgttg agtgtataca acctctatca atctgaagac 360
aagcctgtgg ctctccgtca agctagaaaag attaagagtt tttttgttga ttagtgaatt 420
tttgttttat ttatgtctga tagttcaata aagagacaac acatacaata taaaatcatt 480
gtcttttaaat gttaatttag tagagtgtaa agcctgcatt tttttgtac gcataaacia 540
tgaattcacc ccgcttctgg tttttaaata attatgtcaa actagggaaa attctttttt 600
ttctcttctgt tctttttttg gcttggtgtg gagtcacagg cttgtcttca gattgataga 660
ggttgtatac actcaacaga gcaatcttgg cacgttcgct tcctttttaga tgagctcttg 720
taggatatgt tgtcttctgc tgcttgagta cagatccaat tgcagcatat acttggcaag 780
cagctcgtac accgccttga cagtttgccg gcaacttgct aatgcccttc ttagctctca 840
ccatgatagc gtctgccttt cctacaagct tcagagacaa gcccaacagt ctttgatcac 900
ctaggctacg ggcgttgctt tgctgtatct gttctgtttc ttcttttcta agccattggt 960
gaggcagata acatcgaccc aacatcctcg agccatacta cagcataaaa ggatacgttt 1020
tctttaacag aaatttacct ttttgttatc agcacatata aaaaaaaga aatttaagat 1080
gagtaggact tccattctct caaaaatttt attcaatcca taaatgaatt atttttggac 1140
aaaaaagaaa gattatgcct gattttctct attttttttt tttttacaac tccaccaata 1200
ctttctag 1208

<210> 75

<211> 6316

<212> DNA

<213> *Blakeslea trispora*

<220>

<221> misc_feature

<222> (2694)..(2694)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (4263)..(4263)

<223> n is a, c, g, or t

<400> 75

aaggatgaag aatccaactc taataaaaat cttatggata tctttgatcg actcaaaaag	60
gctttcaatg ctattgctat taataaaaaa gagagagaga gaactatgag caaaaggact	120
ctatgccaaag atggcaaaaa ggcaccagaa acccttagtt tattattgca taatccagtc	180
gagctagtagc ttctgtagct caagcttaac cgaggatcctt ggaatcaact cgtctcgtca	240
ctcttgccga tgatcctaga aatggatatct atggatgtta tactaacatt gttatctttc	300
aaggcctcga agatgttatt gttgcggtga taaataggct gctatgtact gaagttgctc	360
tgtaaaatga atctagttca ctgcctactc agcaaagggt tgtttctaata gtcttttaaag	420
aaagaaaaaa agatacatat agactaccct tcctttcaag actgtaatcg agaatcggcc	480
gatggtttat tacaattaga cgctgggaat aagcaaaagg attcatcttt gtaaataaga	540
gactggtgca tatgaaagca aggatcgtat caaggaatag ttttgatcga gcatcaccag	600
caaatgctgc taatgttggc ttcttctttg cttcctgaga ttgaatggga tgtgcctaga	660
gcattgctat ttttaagtgt atactttaga tttgtgtcct tagatttggtg tcattttatt	720
tagtcaagaa agatccccct ttctctatgt atgctaagaa gaaggagcaa gaagtgtatt	780
tacaagttgg aatgagattg aaatattgta cataataata ataaaaagaa aggtagatca	840
aaaaaatgt tctgcctatt gtaagaaatc gggaccaaca ggtgcttgat aaccagaagt	900

agcttccaat tcaggtagag gctctagga caaatacaca attatgacag gaattttctt 960

gttgacttga acactacaag agaaacgggt cagcacaaaa tccgaaaaaa aaaagaaacg 1020

gaccattcat gtcttaccta tctagctctt tgtcttcaat tgcattccat tgctcaacca 1080

cagatacgct tcccaattga gtatatgat gaagtgttcc ctgcattttt cgcttgacta 1140

attccactac agtcacagtc ttattaatgt tttgtccttt accagtcagg ataatatgat 1200

ctttttgctt cttctatcaa aaaaataatt cttgttttga ataaaaaaaa caaatattta 1260

aagaaactac tttgatgacg gtacctggaa taactcgaga cacacatcta catatgcgtt 1320

gattttattg tggctaattc gaacctcatt ttctgctggg gggggctggt gactttcagt 1380

tgctgagacg tccttcttgc ttcttttata gtcttccact atgattttta tcaagaaagt 1440

aagtcagtga tgattgttac aagctatata tcttgaaaaa gaacagagag gtattattat 1500

cagatgcaac atgggtttct gtatcatttt catttcagtt tctctgttca aaaaaaaaaa 1560

gaacactttc tctttccact cctcaaattt tttctgctaa actcctcgca aaacatgtat 1620

ttgctttaaa ctacaagttg caattgtctg atttagcaat ttcaatatgc cttttgtgaa 1680

tccacccaaa aataaacaag tgcttgagta tacttggggt cagttcaaaa gaaagcaagc 1740

tttttttttt ctttcttggg aaagaaaaaa aaatattggt gagccatcct ttaccagcag 1800

tatgcgagct acgacatagc tgggtctaaca atgactgcaa gcaatagatc gagcttagtc 1860

tttctattgc ttcyttgttt gatctatggt cggccttacg ctgacctatc caatactcga 1920

gataggcaac aagatttcga acagtaatga aataaatttc ggataacagt tgtggatgag 1980

gaagagaaaag cgacttgaac tcgagaaaact ttgttgaaat gaaatccgac cttttacgtg 2040

atcatcatgt attatcctct ttttcttttt tttcgtagtg aattacttac tgattgcgct 2100

caagtcgcgt ctttataaag aagaaaaaaa aatattagaa ctttcaaaaa atataactga 2160

aaataaaagt gtggctcgga gagcaaatac cacatccttt gtcttcgctt tggtaacacg 2220

gttaataaagc cactataggt gaataatgat catttctgag aataaagcgc ggcttgaagc 2280

ttatatccat atcaggattc atattaggca caactcaciaa ttgaggttcc agaagtgcc 2340

atTTTTTTTT cctgatagcc tgtccaatta agatcaaaaa cactgagtt ttctctatat 2400

atTTTTTTTT ttcataattc ttaactcttc ttctctcttc tctctctctc tctctttttg 2460

gcttgcaaaa aaaatcttta gtaataccaa agaaagcaaa ccttttcctt ttcttatttc 2520

cttgcttggt ttttaatttt tgatttctct atgctttaa taccatttc tttctttctt 2580

ctgctattac ctatcttttc attcctctcc cccctctctc tcttggtcta taaacatcat 2640

gaagtcctct tttaaaagtt cgcttgacat ttatgctggt tatatacagc atcntgtggt 2700

ttccaagtgg ttcattcttg cttttgttct ttcgatttcc ctcaacactt atctactgaa 2760

cgcttcgaag caacagccca aagtataat caaaaagggt attgagcggg tagaagtacc 2820

aagtagagaa caacctaaat cagtcataaa gccctcctcc aagaaacact cttctcatca 2880

tcagtctgat gtcattcgcc ctcttgatga agtattgggt ttgctcgga cacccgaggc 2940

cttgactgat gaagagatca tctctattgt tcaagctgggt aaaatggccc cctatgctct 3000

tgaaaaggtc ttgggcgatt tagagcgcgc tgtccatctc cgtcgtgctt tgatctcccg 3060

tgactctcgt acgaaaactt tggaagacag tatgcttccc gtgaaaaact atcattatga 3120

taaagtcatg ggtgcttggt gtgaaaatgt cattgggttat atgcctattc cagtaggtgt 3180

cgcaggtaag aagttcaaca agtcgcgata ttgacaagt tgctcatcat tttcgaaaca 3240

ggtcctttgg tgattgatgg tgattctatt catattccca tggcaactac ggaagggtgt 3300

ttagttgctt ctactgccag aggttgtaaa gcaatcaatg ctgggtgggtg tgccaacaca 3360

attgttggtg ctgatgggtat gactcgaggt ccttggtgctg aatttcctac aatcactcgc 3420

gctgctgact gtaaacgatg gattgaacaa gaggggtgaag ctatcgtgac cgaggcattc 3480
aattcaactt ctcgttttgc tcgtgttcgt aaattgaaag ttgctcttgc cggctcgtcta 3540
gtctacatcc gtttctctac cactacaggt gatgcaatgg gcatgaacat gatctccaag 3600
ggttgtgaaa aggctttaag caagattgct gagagatata ctgatatgca gatcatttct 3660
ctttctggta actattgtac tgacaagaaa cctgctgcta tcaactggat tgaaggacgt 3720
ggtaaactctg ttgttgctga sgctgtcatc cctgggtacgg ttgtcgaaaa ggtattgaag 3780
acctctgtta gtgcttttgg tgagctgaac atctctaaaa acctggttgg ttctgctatg 3840
gctggctccg tcggtggctt taacgctcat gctgctaata ttctaactgc catttacctt 3900
gctactgggc aagatcctgc tcaaaatgta sagagttcta actgtattac tttgatgaaa 3960
gctgtcaatg gcgaaagaga ccttcatata tcttgtacaa tgccctgtat tgaagtaggc 4020
accattgggtg gtggtactat tttgcctcct caacaagcca tgttggattt cattgggtgtg 4080
cgtggtcctc acctaccga acctgggtgcc aatgcccgcw gccttgctcg tgttatctgt 4140
gcctctgtga tggctgggtga attgtcttta tgtgcagctt tggctgctgg tcatcttgta 4200
aaggcacaca tggctcataa tcgtaatacc actgctgctg ccgctgttgt tccctgccct 4260
aanggcatag ttgatgtctc tacacctcct gctacacctg cagaaaagaa tgatcctatt 4320
cctggaagtt gtatcaagtc atagaattaa tattatatat atatcatata caaaaaaag 4380
aaaaaaaaa cactacatct atttatatat ctccatgtac acacacacac acacatataa 4440
aaactcttta ttttccaata ttttgctttt ataaataatc ttatttcatt ctaaataaac 4500
tgtttttttt tattaatcat caaacctgc tgagagctgt gcaatatcat ctatgttttc 4560
atggtttaac tctggtatcg gwcgagcctc ctctgtactt gaagtttgta ggcagttttt 4620
atttaaggct gctggctgat catgatcatc akcaaacctg acagcatgaa gttttgactg 4680

atgagcaatt tcactaaggg cagaatctga actctttcgc ttcctactat tgaccatatt 4740
gtcttttaggt ggaatgagtg aatagcgtct tgtcatatgt aacacagaat caacaatatc 4800
ctggtgatga aactcggcca aacatagcgc ctttctcccc caacaattat aataatcaaa 4860
atgagaatga catgtacggg tttcctcgat gacaatatcc aacgtcttgt cataatcctc 4920
tgtgcygata ccattcatct tttggaagaa cgcacggtag ctctcacaag ctgtcctcag 4980
agagttccgt gccatgtttc ccaatgctcc tggcaagtcg aaatgaagtt gtcgaatctg 5040
gcatgtatg tctacaatgt cgcctgtttc tttcattaga tcaagcattc gtgtagccca 5100
aatgatgtct atgttatgat tttctttcat tccagtaata actatagttt ctcggaat 5160
cgaatgastg atggagtaaa ttcatacaaaa gtgcaagtaa tacatacagt gcttgaagaa 5220
atcttgtgta gcacgcctat attatgtaat ataggatcga ttctcgaaac tcgacataac 5280
caccaggctt tagcaagcgt tttatttcat tcatgacaag ctattgttaa ttcygtgcta 5340
ataaaacaaa atgaaaaaaaa cataccccc tcmaaactta cttcccactc ttgattggaa 5400
aaacaggat agacgtgacg catatgtata taatcaaaac actcatcagg atagggtaaa 5460
ccattgagca catcgcatg ggtgaagaaa gtattaggag gcttgatggc tgtaggatat 5520
atagggtcaa tatcaatacc gtaaaactca gcatttggga attctgtagc catctccaga 5580
atccaagtac ctgtgccaca agcaacatca agcactttag gtaagggtat acattgttgt 5640
tcttgttgtt gttgttgaca atcacttgag tctgagtttc gttttgattg ttttaatgac 5700
aataattctt ttacagggtgc tgagaaatta cgtcaaata gatacttgta aataaaatgc 5760
taaaaataaa aacaatagaa aaaaaattg acgtcattt cattactatg gaaataactg 5820
caaaatctta ccacttgtag aagtctatct tgctcaatct catcgtttgg cagaatgtat 5880
ttattgttgt agtattgata tcttctacca ttcatgatat aactgtcgtc tctaattgctc 5940

tgaggtgaag tacttgtagg tgaaggtgga agtgacgcaa ttttgtcaag cttaacagga 6000
tcctctcggc tacatgtttt ctgcatatca ggaaaatctt gtttatttga aacatcaaca 6060
gtagatgtgg tgtgatcttt ttgaaaata tcgatgcctt cctttgaaag ccttttgaaa 6120
ggctcttttta acttttttga gtgagagcta cccatgatag cttatgaaga attaaaaaga 6180
aaaaagcaaa aaaaattaaa aaaaaaaaaa gtagcaaaaa attctgtcgt aattatacaa 6240
gccaatcaaa atcgaaattc atgcaaggca tagatgttca cgtggatttg atggttgatc 6300
cttttttttt gcaaga 6316

<210> 76

<211> 1170

<212> DNA

<213> *Thermus thermophilus*

<400> 76

atgaagcgcc tttccctgag ggaggcctgg ccctacctga aagacctcca gcaagatccc 60
ctcgccgtcc tgctggcgtg gggccgggccc ccccccggc tcttccttcc cctgccccgc 120
ttccccctgg ccctgatctt tgaccccgag ggggtggagg gggcgctcct cgccgagggg 180
accaccaagg ccaccttcca gtaccgggccc ctctcccgcc tcacggggag gggcctcctc 240
accgactggg gggaaagctg gaaggaggcg cgcaaggccc tcaaagacct cttcctgccg 300
aagaacgtcc gcggctaccg ggaggccatg gaggaggagg cccgggcctt cttcggggag 360
tggcgggggg aggagcggga cctggaccac gagatgctcg ccctctccct gcgcctcctc 420
gggcgggccc tcttcgggaa gcccctctcc ccaagcctcg cggagcacgc ccttaaggcc 480
ctggaccgga tcatggccca gaccaggagc cccctggccc tcttggaact ggccgccgaa 540
gcccgttcc ggaaggaccg gggggcctc taccgcgagg cggaagccct catcgccac 600

352/357

ccgccccctct cccaccttcc ccgagagcgc gccctgagcg aggccgtgac cctcctggtg 660
gcggggccacg agacggtggc gagcgccctc acctggtcct ttctcctcct ctcccaccgc 720
ccggactggc agaagcgggt ggccgagagc gaggaggcgg ccctcgccgc cttccaggag 780
gccctgaggc tctaccccc cgcctggatc ctcaccgga ggctggaaag gcccctcctc 840
ctgggagagg accggctccc cccgggcacc acctggtcc tctcccccta cgtgaccag 900
aggctccact tccccgatgg ggaggccttc cggcccgagc gcttcctgga ggaaagggg 960
acccttcgg ggcgctactt cccctttggc ctggggcaga ggctctgcct ggggcgggac 1020
ttcgccctcc tcgagggccc catcgtcctc agggccttct tccgccgctt ccgcctagac 1080
cccctcccct tccccgggt cctcgcccag gtcaccctga ggcccgaagg cgggcttccc 1140
gcgcggccta gggaggaggt gcgggcgtga 1170

<210> 77

<211> 2981

<212> DNA

<213> Blakeslea trispora

<400> 77

tctagaattc attccattcg aaaggatcaa cataaccaat ttaatgacta ctagctaattg 60
gatacaaata tacgcacaaa aaaagaaaga attctatgat caaagagaac acagacacag 120
agtgatacat ttaaattggtt aagttcttat gatgttaaaa tggtaacttt attattgaat 180
taaatgcgaa tatcgttgct gctttgtact tggaaaacgt taggtaaaag ttggttaatg 240
aaagaagcag gagttgtagt atcatctctt gggaagaaat agaaaaagag gaaagtaaca 300
aagtaacaag caagacaata atagatccaa tggctttcgg tcttacgagt ttgttcagga 360
gcatacttct tttggctatc ttgtaacttt cttggtaagg gattctggcc aaagctttta 420
cagacttggc cggaagtaag cttacttcca gcaagaacga taggaacacc agtacctgga 480

tgtgtactac aaagaaaaga gaaatgagta cgtgcgttat taaaaaaaag aaaaaaagag	540
ggcaaaagta ttacctagct ccgacaaaga aaagattatc ataacggttt gtggaatcct	600
tggtactagg tctgaaccag agaacttgga acacatcatg agaaagacca agaatagaac	660
ctctccaaag gttaaacttg ctttgccaaa cactaggatc attcacttct tcatgttcaa	720
tcaaattagc aaagttgttt actcccaaac gacgttcgat aacttccaga accatcttgc	780
gtgcacgggtt taccaactca ggataatfff cttcagcact gtttcctgtc ttactcttca	840
tatggccaat tggaaccaac acaataatgg agtccttggt gggagggtgcg gcagattcat	900
caattcgaga tggaacgttg acatagaatg aagcttcaga gggcaaaccg aagtcgttga	960
aatctcatc aaaactttcc ttgtaggctt cagccaagaa gatatttgtt acgtctaatt	1020
gaggcacctt tgttgacatg gaccaataaa acgaaataga tgatgaagtg agtttctttg	1080
aggctaattg cttctttgtc caattgcaag gaggtaacag atggtgataa gcataaacia	1140
gatccgcatt acatacgact gcacgggctt caatgacttc tccgctttcc aaagtgcac	1200
cggttacacg cttgtcttta tcgacagtgt taattttagc aacaggcgat tgatatctga	1260
attcagcacc gtactttttg gaggcgatag actcaagctt ctgaacaacc atgttgaaac	1320
caccacgagg ataccagata ccttcagcaa actcgggtgta ttgtaacaaa ctgtaaaactg	1380
ctggagcatc ataaggcgac atactatatt ccaaaaatag aaaatagaac aatgaatatc	1440
aaaattcctt tcaattgccc tttttcacat ttctcttttc ccacccccga ccggtctcac	1500
tcatTTTTTTT tcatccac accacgcgtt gtatgtgtac ttaccccata tacattgttt	1560
gaaaagtaaa agccatacgc attttcttgg tttggaaata tttactggct cggatcataga	1620
tcttaccaaa caagtgcag cgaaagattt caggcacata ctgaagacga atcaaattccc	1680
aatgggtttc aaagttgcgc ttgatagcaa taaatgtacc ttgttcataa tggacatgtg	1740

tttccttcat gaaatccaag aatctaccaa atccaagggg accctcaata cggccaatt 1800
cgcccttcat cttgggttaa tcggaagaga gttgtacggc atcaccgtcg tcaaatgaa 1860
ccttatagtt attgtcacag cgaagcaa ccaatgat accaatacgt tcatccaaat 1920
cagcaaatgc atcttcaaaa agcttaggca tcaaatagag tgagggaccc tgatcaaagc 1980
gatgaccatc gtgatgaatg aatgaacaac ggccaccgga aaagtcgttc ttttcaacaa 2040
cagtaactcg aaaaccttca cgagcaagac gaggcagcagt agcagttccg ccaataccgg 2100
caccaatgac aacaatatgc ttcttttgat cagacatgag attaaaatag ataaggaaaa 2160
gaaagtgaaa agaaattcgg aagcatggca cattcttctt tttataaata catgcctgac 2220
tttctttttc catcgatatg atatatgcat atgatagata tacaagcaat cttcttcaag 2280
gagtttgaaa tttgtcctc caggagcaaa aaaaagtttt tttttataca tgtttgtaca 2340
caagaatagt taccaatttg ctttgggtctt acgtgctgca agtttatatc gttttcaatt 2400
tctttgtctt tacattttct ttgtccttta tctttcctca ttagtcttt gggagaatta 2460
ggaaaaggga gcggaaagg aagaaatgct tgcgtatttt actaattcgg caaacatcca 2520
atttggcaaa cagcagcctg tgcaacgctc tcgagatgac agtatctttg attacactct 2580
aatctctgat gacccgacca aaaagagcga acaagaaat aatcttgtgc attcgaatat 2640
gatggaagat tttttccccc ttattctaaa tgttgacata gcgtgtatgt tatataaaca 2700
aaaagaaatt gtacaaactt tcttttcttc tctttttatt ttatctctat gtcaatactc 2760
acttatctgg aatttcatct ctactataca ctacctgtcc ttgcggcatt gtgttggtg 2820
ctaaagccgt ttcactcaca gcaagacaat ctcaagtata aatttttaat gttgatggcc 2880
gcctctaccg catogatttg ggacaattat atcgtttate atcgcgcttg gtggtactgt 2940
cctacttggtg ttgtggctgt cattggctat gtacctctag a 2981

<210> 78

<211> 1749

<212> DNA

<213> *Blakeslea trispora*

<400> 78

atgtctgac	aaaagaagca	tattgttgtc	attggtgccg	gtattggcgg	aactgctact	60
gctgctcgtc	ttgctcgtga	aggttttcga	gttactgttg	ttgaaaagaa	cgacttttcc	120
ggtggccggtt	gttcattcat	tcatcacgat	ggtcacgcgt	ttgatcaggg	tcctcactc	180
tatttgatgc	ctaagctttt	tgaagatgca	tttgctgatt	tggatgaacg	tattggtgat	240
catttggtatt	tgcttcgctg	tgacaataac	tataagggtc	attttgacga	cggtgatgcc	300
gtacaactct	cttccgattt	aaccaagatg	aagggcgaat	tggaccgtat	tgaggggtccc	360
cttggatttg	gtagattcct	ggatttcacg	aaggaaacac	atgtccatta	tgaacaaggt	420
acatttattg	ctatcaagcg	caactttgaa	accatttggg	atttgattcg	tcttcagtat	480
gtgcctgaaa	tctttcgctt	gcacttggtt	ggtaagatct	atgaccgagc	cagtaaatat	540
ttccaaacca	agaaaatgcg	tatggctttt	acttttcaaa	caatgtatat	gggtatgtcg	600
ccttatgatg	ctccagcagt	ttacagtttg	ttacaataca	ccgagtttgc	tgaaggtatc	660
tggtatcctc	gtggtgggtt	caacatgggt	gttcagaagc	ttgagtctat	cgcctccaaa	720
aagtacggtg	ctgaattcag	atatcaatcg	cctgttgcta	aaattaacac	tgctgataaa	780
gacaagcgtg	taaccggtgt	cactttggaa	agcggagaag	tcattgaagc	cgatgcagtc	840
gtatgtaatg	cggatcttgt	ttatgcttat	caccatctgt	tacctccttg	caattggaca	900
aagaagacat	tagcctcaaa	gaaactcact	tcatcatcta	tttcgtttta	ttggtccatg	960
tcaacaaagg	tgcttcaatt	agacgtacac	aatatcttct	tggctgaagc	ctacaaggaa	1020

356/357

agttttgatg agattttcaa cgacttcggt ttgccctctg aagcttcatt ctatgtcaac 1080
gttccatctc gaattgatga atctgccgca cctcccaaca aggactccat tattgtgttg 1140
gttccaattg gccatatgaa gagtaagaca ggaaacagtg ctgaagaaaa ttatcctgag 1200
ttggtaaacc gtgcacgcaa gatggttctg gaagttatcg aacgtcgttt gggagtaaac 1260
aactttgcta atttgattga acatgaagaa gtgaatgac ctagtgtttg gcaaagcaag 1320
tttaaccttt ggagagggtc tattcttggt ctttctcatg atgtgttcca agttctctgg 1380
ttcagacctc gtaccaagga ttccacaaac cgttatgata atcttttctt tgcggagct 1440
agtacacatc caggtactgg tgctcctatc gttcttgctg gaagtaagct tacttccgac 1500
caagtctgta aaagctttgg ccagaatccc ttaccaagaa agttacaaga tagccaaaag 1560
aagtatgctc ctgaacaaac tcgtaagacc gaaagccatt ggatctatta ttgtcttgct 1620
tgttactttg ttactttcct ctttttctat ttcttcccaa gagatgatac tacaactcct 1680
gcttctttca ttaaccaact ttacctaac gttttccaag taaaagcag caacgatatt 1740
cgcatthaa 1749

<210> 79

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 79

ccgatggcga cgacggaagg ttgtt

25

<210> 80

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 80

catgttcattg cccattgcat cacct

25